

GEOS-5 seasonal forecast system

Anna Borovikov^{1,2,*}

Richard Cullather^{1,3}

Robin Kovach^{1,2}

Jelena Marshak¹

Guillaume Vernieres^{1,2}

Yury Vikhliakov^{1,4}

Bin Zhao^{1,5}

Zhao Li^{1,2}

¹Global Modeling and Assimilation Office, National Aeronautics and Space Administration
Goddard Space Flight Center, Greenbelt, MD, USA

²Science Systems and Applications Inc, Lanham, MD, USA

³Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD,
USA

⁴Universities Space Research Association, Columbia, MD, USA

⁵Science Applications International Corporation, McLean, VA, USA

*Anna.Borovikov@nasa.gov, phone: 301-614-6030, fax: 301-614-6246

Abstract

Ensembles of numerical forecasts based on perturbed initial conditions have long been used to improve estimates of both weather and climate forecasts. The Goddard Earth Observing System (GEOS) Atmosphere-Ocean General Circulation Model, Version 5 (GEOS-5 AOGCM) Seasonal-to-Interannual Forecast System has been used routinely by the GMAO since 2008, the current version since 2012. A coupled reanalysis starting in 1980 provides the initial conditions for the 9-month experimental forecasts. Once a month, sea surface temperature from a suite of 11 ensemble forecasts is contributed to the North American Multi-Model Ensemble (NMME) consensus project, which compares and distributes seasonal forecasts of ENSO events. Since June 2013, GEOS-5 forecasts of the Arctic sea-ice distribution were provided to the Sea-Ice Outlook project. The seasonal forecast output data includes surface fields, atmospheric and ocean fields, as well as sea ice thickness and area, and soil moisture variables. The current paper aims to document the characteristics of the GEOS-5 seasonal forecast system and to highlight forecast biases and skills of selected variables (sea surface temperature, air temperature at 2 m, precipitation and sea ice extent) to be used as a benchmark for the future GMAO seasonal forecast systems and to facilitate comparison with other global seasonal forecast systems.

Keywords: global forecast, seasonal prediction, NMME, GEOS-5, ENSO, forecast skill

1 Introduction

Deterministic numerical weather prediction forecasts have a forecasting window that is limited to about 15 days (e.g., Lorenz, 1963, 1993). As noted by Palmer and Anderson (1993, 1994) and others, useful predictability is possible beyond this limit in part because boundary forcing such as sea surface temperatures or soil moisture (Koster and Suarez, 2001) may vary slowly and reliably, and may then influence statistics of the atmosphere. In 2010, the US National Academies reported on the state of seasonal-to-interannual predictability, and suggested avenues for progress (Weller et al., 2010). Among the recommendations was the need to establish and evaluate a multi-model ensemble, which was recognized as a viable strategy for resolving forecast uncertainty (e.g., Kirtman et al., 2014). The NASA Global Modeling and Assimilation Office (GMAO) has participated in the North American Multimodel Ensemble (NMME; Kirtman et al., 2014) since its inception. The purpose of the NMME is to advance the capabilities of the climate prediction models, and utilize the system in a near-operational mode to demonstrate feasibility. The GMAO system is based on its use and experience with data assimilation methods that have been developed for mission support and to enhance NASA's program of earth observations. The development and use of the seasonal forecasting system enhances the use of NASA data and contributes to observing system science by improving assimilation systems and atmosphere and ocean modeling tools. Evaluation of the GMAO system has previously been conducted with a focus on the predictability of the El Niño/Southern Oscillation phenomenon (ENSO; Ham et al., 2014; Vernieres et al., 2012). In this paper, we provide a more comprehensive assessment of current forecasting system as it reaches the end of its life cycle. As the system has now progressed through several years within the NMME near-operational mode, this paper critically examines recent performance.

The layout of the paper is as follows. Section 2 provides an overview of the GMAO Goddard Earth Observing System (GEOS) Atmosphere-Ocean General Circulation Model, version 5 (GEOS-5 AOGCM) Seasonal Forecast System. Section 3 details the initialization procedure for each system component, and the means of ensemble generation through field perturbations and sampling in time. Section 4 presents an assessment of the forecast sea surface temperature (SST). Section 5 presents the bias and skills of the relevant atmospheric fields, including 2-m air temperature (T2M) and precipitation. Section 6 examines the prognostic sea ice cover. Conclusions are presented in section 7.

2 Overview of the GEOS-5 Seasonal Forecast System: Model Components

The GEOS-5 AOGCM has been developed to simulate climate variability on a wide range of time scales, from synoptic time scales to multi-century climate change, and has been tested in coupled simulations and in data assimilation mode. The ocean and atmosphere exchange fluxes of momentum, heat and freshwater through a "skin layer" interface which includes parameterization of the diurnal cycle and a sea ice model. All components are coupled together using the Earth System Modeling Framework (ESMF, Hill et al. 2004). The goal in having a multi-scale modeling system with its different components communicating through a unified interface (ESMF) is to be able to propagate improvements made to a physical process in one component to the other components smoothly and efficiently. The GEOS-5 AOGCM was configured to participate in the Coupled Model Intercomparison Project phase 5 (CMIP5), which provides a standard protocol for evaluation of coupled GCMs. To evaluate the model's ability to simulate the Earth's climate, it was validated against observational data and reanalysis products.

2.1 Atmospheric Component

The atmospheric component of the GEOS-5 AOGCM is Fortuna-2.5, the same that was used for the Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011), but with adjusted parameterization of moist processes and turbulence (Molod et al. 2012). The model has a finite volume dynamical core (Lin 2004), which is integrated with various physical packages through the ESMF.

The physics package includes parameterization of moist processes, radiation, turbulent mixing and surface fluxes. The moist component contains parameterization of convection using the Relaxed Arakawa-Schubert scheme (Moorthi and Suarez 1992), and the large-scale precipitation and cloud cover model as described in Baumeister et al.

(2006). The radiation component includes parameterization for long wave (Chou 1990, 1992) and short wave radiation processes (Chou et al 1994). The turbulence component consists of parameterization for vertical diffusivity, the planetary boundary layer and gravity wave drag. The free atmospheric turbulent diffusivities are based on the gradient Richardson number. The parameterization of the boundary layer is based on Lock et al. (2000) scheme, acting together with scheme of Louis and Geleyn (1982). The Lock et al. (2000) scheme includes a representation of non-local mixing (driven by both surface fluxes and cloud-top processes) in unstable layers, either coupled to or decoupled from the surface, and an explicit entrainment parameterization. The original scheme was extended in GEOS-5 to include moist heating and entrainment in the unstable surface parcel calculations. GEOS-5 incorporates two gravity wave drag parameterizations, an orographic gravity wave drag formulation based on McFarlane (1987), and a formulation for non-orographic waves based on Garcia and Boville (1994). The surface exchange of heat, moisture and momentum between the atmosphere and land, ocean or sea ice surfaces are treated with a bulk exchange formulation based on Monin-Obukhov similarity theory. The atmospheric model uses a Cartesian grid with a $1^\circ \times 1.25^\circ$ horizontal resolution and 72 hybrid vertical levels with the upper most level at 0.01 hPa.

2.2 Ocean Component

The ocean component of the GEOS-5 AOGCM is the Modular Ocean Model version 4 (MOM4) developed at Geophysical Fluid Dynamics Laboratory (Griffies 2012). It is a non-Boussinesq, hydrostatic, primitive equations model with a staggered Arakawa B-grid or C-grid and generalized level (vertical) coordinate based on depth or pressure. A tripolar grid is used to resolve the Arctic Ocean without polar filtering (Murray, 1996). The nominal resolution of the ocean grid is $\frac{1}{2}^\circ$, with a meridional equatorial refinement to $\frac{1}{4}^\circ$. It is a regular Cartesian grid south of 65°N , and curvilinear north of 65°N , with two poles located on land to eliminate the problem of vanishing cell area at the geographic North Pole. The resulting tripolar grid has a minimum and maximum resolution of 15 km and 52 km, respectively.. The ocean topography is derived from the ETOPO5 data set (Smith and Sandwell 1997). The topography is represented as a partial bottom step to better represent topographically influenced advective and wave processes. Vertical mixing follows non-local K-profile parameterization of Large et al. (1994) and includes parameterizations of tidal mixing on continental shelves (Lee et al. 2006) as well as breaking internal gravity waves (Simmons et al. 2004). Mesoscale eddy transport uses the method developed by Ferrari et al. (2010), modifying the isoneutral method developed by Gent and McWilliams (1990). The restratification effect of submesoscale eddies uses the theory developed by Fox-Kemper et al. (2008) and implementation by Fox-Kemper et al. (2011). The horizontal viscosity uses the anisotropic scheme of Large et al. (2001) for better representation of equatorial currents. The exchange with marginal sea is parameterized under coarse resolution as discussed in Griffies et al (2004).

2.3 Sea Ice Component

The sea ice component of the GEOS-5 AOGCM is the Community Ice CodE, version 4 (CICE; Bailey et al. 2010, Hunke 2008) developed at Los Alamos National Laboratory. The model includes several interacting components to allow for semi-implicit coupling between the atmosphere and ice surface: a thermodynamic model that computes local growth rates of snow and ice due to vertical conductive, radiative and turbulent fluxes, along with snowfall; a model of ice dynamics, which predicts the velocity field of the ice pack based on a model of the material strength of the ice; a transport model that describes advection of the area concentration, ice volumes and other state variables; and a ridging parameterization that transfers ice among thickness categories based on energetic balances and rates of strain. A skin layer interface is used for the exchange of basal heat, salt, and freshwater fluxes with the underlying MOM4 ocean model; ice pressure is not exerted on the ocean. The CICE model is configured with standard settings but without the use of melt ponds.

2.4 Land Component

The land surface model in the GEOS-5 AOGCM is a catchment-based hydrological model described in Koster et al. (2000). In this model, subgrid heterogeneity in surface moisture state is treated statistically. The applied subgrid scale distributions are related to the topography, which exerts a major control over much of the subgrid variability. The catchment model is coupled to the multi-layer snow model described in Stieglitz et al. (2001).

3 Overview of the GEOS-5 Seasonal Forecast System: Initial State Generation

3.1 Atmosphere Initialization

In the coupled model initialization, selected atmospheric variables are constrained with the Modern-Era Retrospective Analysis for Research and Application (MERRA; Rienecker et al. 2011). These variables include surface pressure, pressure thickness, zonal and meridional winds, specific humidity, ozone concentration, and potential temperature.

3.2 Ocean and Sea-Ice Initialization

The Goddard Earth Observing System integrated Ocean Data Assimilation System (GEOS-iODAS) is used for both ocean state and sea ice initialization for the production of analysis products (MERRA-Ocean). The ocean and sea-ice initialization methodology is described in detail in Vernieres et. al. (2012) and Keppenne et al. (2008). An overview of the initialization procedure relevant to the hindcasts is presented here.

The assimilated observing system consists of:

- sea surface temperature observations from CMIP5 (Hurrell et al. 2008) prior to 1982 and Reynolds (Reynolds et al. 2007) from 1982 to present;
- temperature and salinity profiles from eXpendable Bathythermographs (XBTs) and Conductivity Temperature Depth (CTD) sensors extracted from the EN3 data base (Ingleby and Huddleston 2007) with time-varying XBT corrections applied according to Levitus et al. (2009), the tropical moored buoy array (McPhaden et al. 2010) - TAO/TRITON, PIRATA, and RAMA arrays and Argo floats, with profiles from the Argo Global Data Assembly Center (GDAC);
- sea ice concentration from the National Snow and Ice Data Center (NSIDC).

The NSIDC sea-ice concentrations product is based on passive microwave observations of ice concentration from the Nimbus-7 Scanning Multi-channel Microwave Radiometer (SMMR) and the Defense Meteorological Satellite Program (DMSP) Special Scanning Microwave Imager (SSM/I) and Special Scanning Microwave Imager/Sounder (SSMIS). It has a 25 km spatial resolution for both the north and south polar regions. Temporal resolution is every other day from October 1978 to July 1987 (SMMR), then daily from August 1987 to present (SSM/I, SSMIS). Ice concentrations from CMIP5 and Reynolds are used in areas that are not measured due to orbit inclination (poleward of 87.2° for SSM/I and 84.5° for SMMR).

The above observations are assimilated using an ensemble optimal interpolation technique (Oke et al. 2010, Wan et al. 2010) with 5-day window from 1979 to present. The model is also weakly constrained to the World Ocean Atlas 2009 (WOA09) gridded climatology (Antonov et al. 2010; Locarnini et al. 2010) of $T(z)$ and $S(z)$ at 1° degree resolution and from 0 to 4500m and of Sea surface salinity (SSS) to correct some of the model's biases, particularly prior to the Argo era.

The resulting analysis (MERRA-Ocean) has been extensively diagnosed through The Ocean Reanalyses Intercomparison Project (Balmaseda et al. 2015) in terms of various parameters such as mixed-layer depth, thermocline depth, heat and salinity content, overturning circulation, etc.

3.3 Land

An important aspect of the GEOS-5 initialization concerns the treatment of the land. Observed precipitation data are used to construct a corrected version of the hourly MERRA (or GEOS-5 forward processing) precipitation fields, which are then used to force the land surface and generate enhanced soil moisture initial conditions for initializing the GEOS-5 seasonal forecasts. The corrections to the precipitation are obtained using the Global Precipitation Climatology Project version 2.1(GPCPv2.1, provided by the NASA/Goddard Space Flight Center's Laboratory for Atmospheres, which develops and computes the dataset as a contribution to the GEWEX Global Precipitation Climatology Project) and Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP, provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their web site at <http://www.esrl.noaa.gov/psd/>) pentad precipitation data following Reichle et al. (2011). As the first step, the CMAP dataset is rescaled to match the (seasonally variable) long-term climatology of the GPCP. During the second step, hourly MERRA total precipitation is time averaged and re-gridded to the scale of the correcting CMAP dataset (i.e., to pentad and 2.5° resolution). Next, separately for each pentad of each year and for each 2.5° grid cell, a scaling factor is computed by determining

the ratio of the (climatologically adjusted) CMAP estimate to the MERRA data (i.e., on the grid and at the time scale of the correcting observations). Finally, these scaling factors are re-gridded back to the MERRA grid and a scaling factor derived for a given grid cell and year/pentad is applied to the MERRA precipitation rates (large-scale precipitation, convective precipitation, and snowfall separately) in each of the 120 hourly time steps within that pentad. By construction, the corrected MERRA precipitation is nearly identical to the CMAP estimates at the pentad and 2.5° resolution. The diurnal cycle, the frequency and relative intensity of rainfall events at the sub-pentad scale, and the sub-2.5° spatial variations are entirely based on MERRA estimates.

3.4 Sampling in Time

Each month, the GMAO produces an ensemble of 12 (13 in November) real-time GEOS-5 coupled model forecasts. The ensemble is produced by initializing the model every five days (table 1) prior to the start of the month, except for the date closest to the start of the month when additional GEOS-5 forecasts are generated by various perturbation methods (tables 2 and 3). The perturbations are produced using a breeding approach (perturbing the atmosphere and/or ocean), and a simple scaled differencing approach involving nearby (in time) atmosphere and ocean states. We note that due to time constraints only 11 ensemble members are delivered to the (North American Multi-Model Ensemble (NMME) project (forecasts are due by the 8th of the month).

In addition to the forecasts, the GMAO produced a suite of hindcasts (1982-2012) used to calibrate/bias correct the forecasts and assess forecast skill. The ensemble members for the hindcasts are produced in the same way as for the forecasts and span the period 1982-2012.

Table 1 Seasonal Forecast Schedule

Seasonal Forecast Schedule											
Jan		Feb		Mar		Apr		May		Jun	
12	12	1	11	2	10	3	12	4	11	5	11
12	17	1	16	2	15	3	17	4	16	5	16
12	22	1	21	2	20	3	22	4	21	5	21
12	27	1	26	2	25	3	27	4	26	5	26
1	1	1	31	3	2	4	1	5	1	5	31
1	6	2	5	3	7	4	6	5	6	6	5
Jul		Aug		Sep		Oct		Nov		Dec	
6	10	7	10	8	9	9	8	10	8	11	12
6	15	7	15	8	14	9	13	10	13	11	17
6	20	7	20	8	19	9	18	10	18	11	22
6	25	7	25	8	24	9	23	10	23	11	27
6	30	7	30	8	29	9	28	10	28	12	2
7	5	8	4	9	3	10	3	11	2	12	7
								11	7		

Table 1 shows the start dates for the ensemble members of the GMAO Seasonal forecasts and hindcasts. The bold shaded values denote the closest dates to the start of the month for which additional ensemble members are generated using various perturbation methods.

3.5 Perturbations

Ensembles of numerical forecasts based on perturbed initial conditions have long been used to improve estimates of both weather and climate forecasts. The GEOS-5 seasonal forecast is arranged so that it uses a set of 7 members for the forecast initialized on the day closest to the beginning of the month, and one member otherwise for a combined ensemble of 11 members. Initial perturbation method, that adds small perturbation to analysis initial conditions, is used to generate the ensemble forecast. Since one goal has been the use of the ensemble spread as an indicator of expected forecast skill, bred vectors (Toth and Kalnay, 1993) have been used as perturbations to capture the fastest growing modes on weather time scales. More recently, the coupled breeding method was developed for coupled atmosphere ocean systems to capture the dominant mode of coupled instabilities associated with the El Niño/Southern Oscillation (ENSO) (Cai et al. 2003; Yang et al. 2006, 2008; Ham et al. 2012).

The breeding is applied from 1980 with the aim of capturing the fastest-growing errors in the seasonal forecasts. Two-sided breeding is applied, which means positive and negative bred runs are restarted every month by adding

and subtracting the bred vector to the initial conditions generated from the Ensemble Optimal Interpolation (EnOI) option of the GEOS ocean data assimilation system, forced with NASA's Modern-Era Retrospective analysis for Research and Applications (MERRA) (Rienecker et al. 2011).

The rescaling interval chosen for the breeding is 30 days. The rescaling norm is the RMS difference of the instantaneous sea surface temperatures (SSTs) from the positive and negative bred runs; the region for defining the norm is the tropical Pacific domain over 120°E-90°W, and 10°S-10°N. At every re-initialization during the breeding cycle, perturbations are re-scaled so that the magnitude of the norm is reduced to 10% of the natural variability of SST over the norm region (i.e. 0.48°C). Another method to perturb initial conditions is based on the GEOS-5 analysis on two different days. Similar to breeding, the perturbations are re-scaled and the magnitude of the norm is reduced to 10% of the natural variability of SST over the norm region (i.e. 0.48°C). A combination of these two methods is used in generating the ensemble members for the seasonal forecast.

Tables 2, 3 and 4 illustrate the perturbations of the initial conditions for all ensemble members generated at the beginning of the month that are submitted to NMME. Additional ensemble members utilize satellite altimetry data, which do not cover the full NMME hindcast period. Table 2 Ensemble members' perturbation combinations

Initial conditions (IC) perturbations							
Ensemble member	1	2	3	4	5	6	7
IC type							
Ocean	O	O B-	O B+	O I-	O I+	O	O B+
Atmosphere	A	A B-	A B+	A I-	A I+	A I-	A

Table 3 Format of IC perturbations

Ensemble member	Perturbation type
1	ocean and atmosphere IC are not perturbed
2	ocean and atmosphere IC are perturbed using negative bred vectors
3	ocean and atmosphere IC are perturbed using positive bred vectors
4	ocean and atmosphere IC are perturbed using negative rescaled difference between two analyses
5	ocean and atmosphere IC are perturbed using positive rescaled difference between two analyses
6	atmosphere IC are perturbed using negative rescaled difference between two analyses
7	ocean IC are perturbed using positive bred vectors

Table 4 List of perturbed variables

Perturbed variables	
ocean model grid	temperature and salinity
	ocean velocities
	surface temperature, salinity and velocities
	sea level and frazil
	ice velocity and strain rate components
	ice strength, extent and stress tensor components
atmosphere model grid	wind components
	potential temperature
	surface pressure
	specific humidity
skin layer tiles	skin temperature, salinity and depth

4 Forecast Skill: SST

The forecast accuracy of the coupled model forecasts is assessed by the amplitude and phase of SST anomaly measured for specified regions and by the global patterns of SST. The forecast accuracy of the atmosphere-land forecasts will be assessed by the patterns and amplitude of the precipitation and surface temperature anomalies. For the arctic sea ice forecasts evaluation, the sea ice extent is compared with observations and against other similar systems participating in the Ice Outlook project.

Fig1 depicts the regions that are used to compute the SST indices routinely used to assess the forecast skills. Other regions of interest, used for the case studies of the 2m air temperature and precipitation, are also shown.

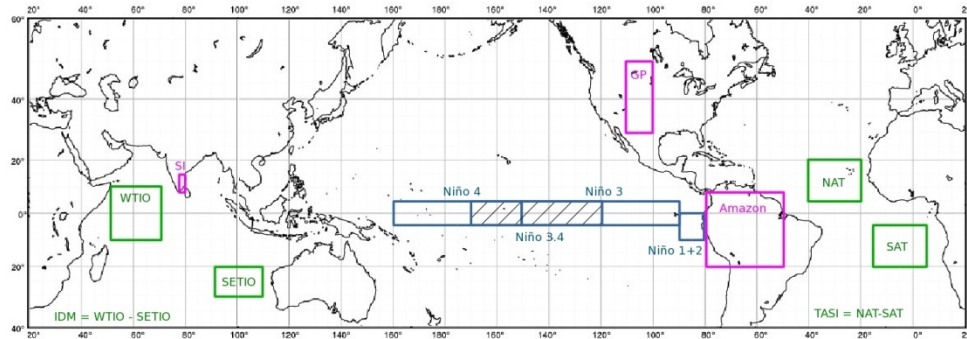


Fig 1Regions used in SST forecast skill assessment are shown by the colored, blue and green, annotated rectangles, Niño 3.4 region is hatched. The three regions shown by pink rectangles are used in the 2m air temperature and precipitation evaluation

In the equatorial Pacific Ocean the easternmost Niño 1+2 region boundaries are 0° - 10° S, 90° W- 80° W, the eastern Niño 3 region boundaries are 5° N- 5° S, 150° W- 90° W, the central Niño 3.4 region boundaries are 5° N- 5° S, 170° W- 120° W and the western Niño 4 region boundaries are 5° N- 5° S, 160° E- 150° W. In the Indian Ocean the Western Tropical Indian Ocean (WTIO) SST anomaly index is calculated in the box 50° E - 70° E, 10° S - 10° N, the Southeastern Tropical Indian Ocean (SETIO) SST anomaly index is calculated in the box 90° E - 110° E, 10° S - 0° ; the Dipole Mode Index (DMI) is calculated as the difference of the WTIO and SETIO indices(Saji et al. 1999). The Tropical Atlantic SST Index (TASI) is defined as the differences between the North Atlantic Tropical (NAT) and South Atlantic Tropical (SAT) SST indices, computed in the boxes 40° W - 20° W, 5° N - 20° N and 15° W - 5° E, 20° S - 5° S respectively (Chang et al. 1997)

4.1 Forecast Drift

Forecast drift is an artifact of the imperfect models. For the seasonal forecast it is necessary to properly account for the drift and calibrate the forecast accordingly. A continuous coupled analysis and a complete set of retrospective forecasts for the entire training period are required to consistently de-trend the forecast. In GEOS-5 system the drift is calculated as the average of these hindcasts from 1981-2010 for every ensemble member. It is subsequently subtracted from the production forecasts. This method of drift removal follows the convention established by Stockdale (1997) and others. The forecast bias characteristics are also important to understand for evaluating the performance of the current and the future seasonal-to-interannual forecast systems. Comparison of the retrospective forecasts to the observations is helpful in determining the model's skill.

4.1.1 Global Bias

Fig 2 shows the global forecast drift from Reynolds SST Climatology for December. Nine panels (top to bottom, left to right) correspond to initial conditions 1 month prior to December, 2 months prior and so on, the last panel shows the forecast for December initialized in April. This is the average drift of all the ensemble members. Immediately one can see from the first panel the cold bias appearing during the first month of the forecast in the northwestern Atlantic ocean where subpolar surface water displaces the warm, salty water of the North Atlantic Current(Large and Danabasoglu 2006) and off the east coast of South America at the confluence of the Brazil current and the Antarctic Circumpolar Current exiting Drake Passage. Just as quickly the warm biases develop in the coastal areas off the west coasts of South America and southern Africa and off the east coasts of Asia and North America. The biases described above are present in all the forecast regardless of the initialization time. The coupled model exhibits a cooling trend in the equatorial Pacific Ocean and in the Southern Ocean, but it takes about 4 to 5 months for the

large scale SST biases in these regions to reach their maximum value of about 3°C (see panels in the middle row of Fig 2, the December forecasts initialized in August and July). There is a cold bias in the southeastern tropical part of every of the three major ocean basins (Indian, Pacific, Atlantic) developing over the same time period.

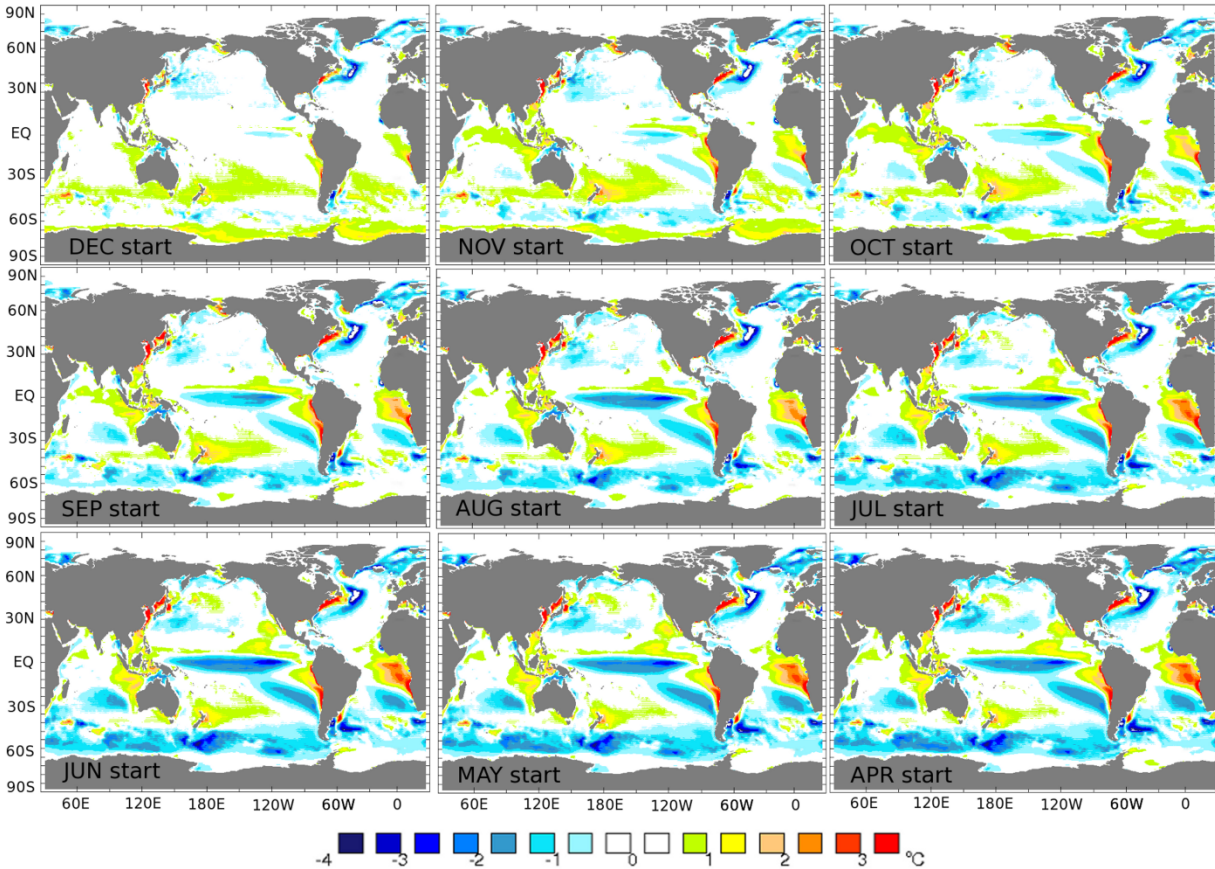


Fig 2 Monthly mean SST forecast drift with respect to Reynolds climatology for December for every forecast lead time, i.e. December forecast initialized at the beginning of December (lead 1), at the beginning of November (lead 2), up to the December forecast initialized in April (lead 9)

Fig 3 shows the model drift for every forecast for the last (lead 9) month, when the amplitude of the drift is expected to be the largest. The top left panel on this figure would contain the lead 9 drift of the December forecast (initialized in April, shown on the bottom-right panel of the Fig 2, thus omitted here). The order of panels is schematically listed in its place: predicted month first and next to it in parenthesis the initialization month. Thus the top row shows the model bias for winter (initialized in Apr-Jun), the second row shows the bias for spring (initialized in Jul-Sep), etc. The model climatology is colder than Reynolds along the equator in the Pacific Ocean during all seasons, but especially so in the fall and winter (initialized in Jan-Mar). The southeastern tropical Pacific and southeastern tropical Atlantic cold biases are also present throughout the year, but more pronounced in the boreal winter (Dec-Mar) season (initialized in Mar-Jul). In the Indian Ocean, the cold bias in the tropical southeastern part and along the equator and concurrent warm bias off the western Australia coast is present only during Dec-Mar (initialized in Apr-Jul).

The SST bias in the northern Pacific Ocean has the strongest seasonality: there is a dipole structure with warm mid latitudes and cold tropics in Jul-Sep (initialized in Nov-Jan) with the differences between the model and observed climatologies as large as +3°C and -2°C. At the same time a similar pattern of warm mid latitudes/cold tropics bias appears in the Atlantic ocean, but the magnitude of the bias is smaller, with warm bias about 1°C and cold about

295 2°C.

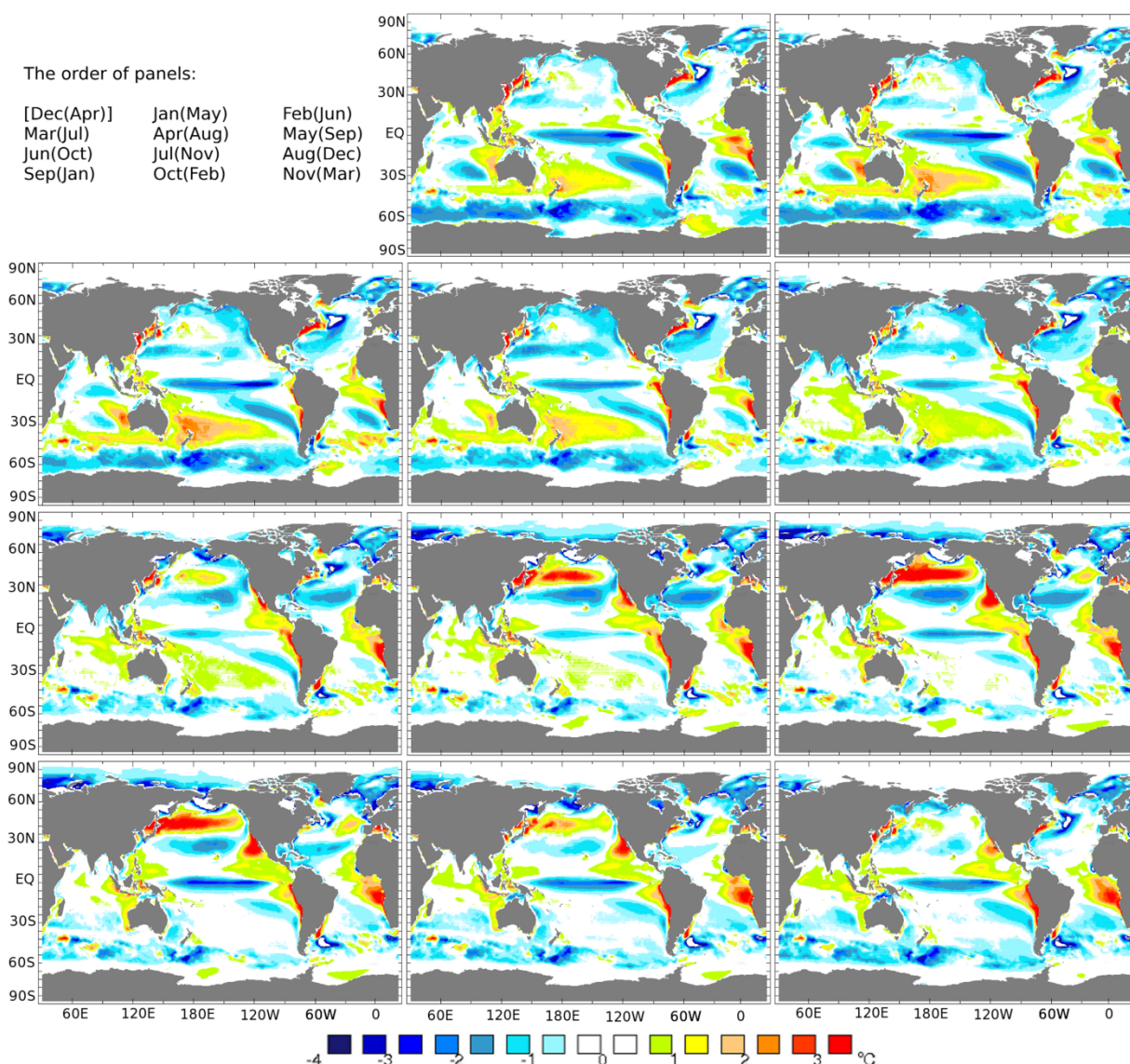


Fig 3Lead 9 monthly mean Forecast SST forecast bias with respect to Reynolds climatology for Jan-Nov predicted months; the month when the forecast was initialized is shown in parenthesis; lead 9 shown (the order of panels is shown in the place of Dec(Apr) predicted(initialized) month)

4.1.2 ENSO, IDM and TASI SST indices

Fig 4 shows the mean ensemble bias for the Pacific Ocean Niño 3, Niño 4, Niño 3.4 and Niño 1+2 indices, WTIO, SETIO, IDM and TASI computed with respect to Reynolds SST. In the western and central equatorial Pacific Ocean (Niño 4, Niño 3.4) the model has an exaggerated seasonal cycle with a cold bias of up to 2°C in the boreal fall and winter for all forecasts targeting this time period. The fastest drift away from the observations appears in the forecasts initialized in summer (July, August). The forecasts initialized in winter and early spring (Jan-Mar) tend to stay close to the observations until the onset of summer. In the easternmost Pacific Ocean (Niño 1+2 region) the model is biased warm up to 2°C throughout the year with the exception of winter target months, when all forecasts return close to the observations. The Niño 3 region has the smallest bias and the most accurate seasonal cycle represented by the model. In the Indian Ocean the model in general is less biased: the warm bias is slightly larger in the east than in the west, thus the IDM (Indian Ocean dipole mode index) is slightly biased towards negative values during the summer and fall. The Tropical Atlantic SST Index (TASI)– the difference between the northern and the

southern Atlantic ocean control regions (refer to Fig 1 for their definition) – is negatively biased by approximately 1°C throughout the year, underestimating the absolute value of the gradient between the north and the south index poles in the summer and fall, and overestimating it in the spring.

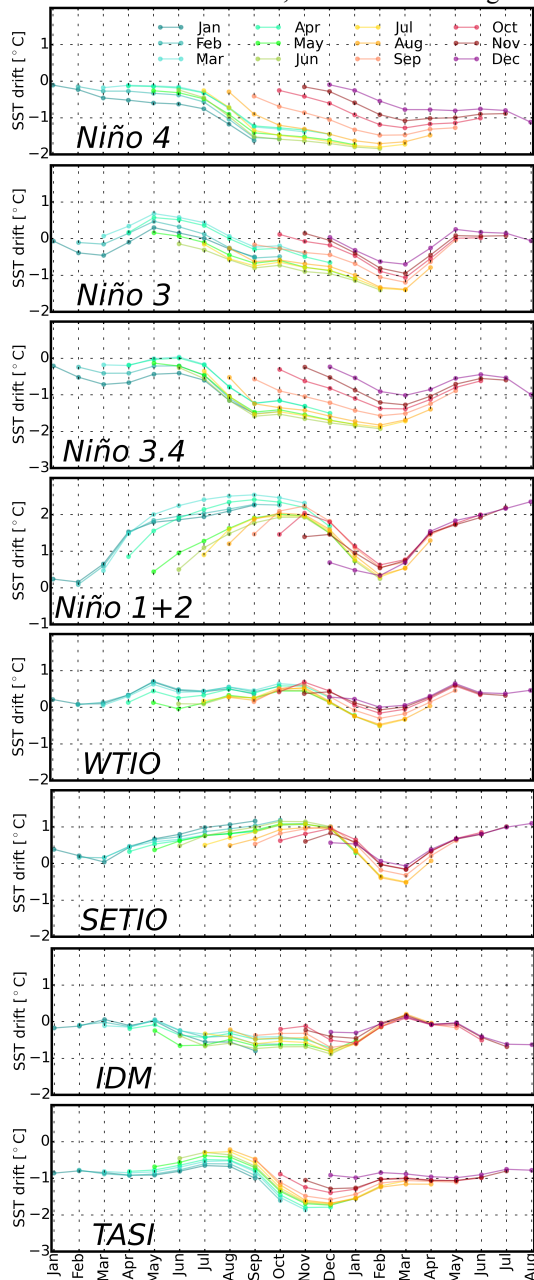


Fig 4Monthly mean SST forecast drift with respect to Reynolds for equatorial Pacific, Indian and Atlantic Ocean indices

4.2 Forecast Skill

Similarly to the forecast drift discussion, the SST global skill maps are presented first, followed by the analysis of the regional indices. Anomaly Correlation Coefficient (ACC) is used as a measure of potential skill and Mean Square Skill Score (MSSS) as a measure of actual skill. MSSS is computed with respect to climatology, i.e. zero anomaly case, as follows,

$$MSSS_{clim} = \frac{MSE_{clim} - MSE_{fcst}}{MSE_{clim}}, \text{ where } MSE_{fcst} = \frac{1}{n} \sum_{i=1}^n \left(T_{fcst}(i) - T_{obs}(i) \right)^2$$

Here $T_{fcst}(i)$ is the temperature anomaly of the i -th hindcast and $T_{clim}(i) \equiv 0$.

4.2.1 Global Anomaly Correlation Skill

Fig 5 shows the global SST ACC computed for all forecast from all initializations months combined, with each panel representing the lead months. Lead month 1 has high correlation (above 0.8) with Reynolds SST in all the ocean basins. By lead months 2 and 3, the high correlation remains only in the tropical Pacific and Atlantic oceans. The Atlantic Ocean skill drops below 0.6 by lead month 6, but still remains high in the north Atlantic Iceland Basin region. Only in a portion of the Equatorial Pacific (Niño 3.4 region) ACC remains above 0.6 by lead month 9. From the significance point of view, the skill across equatorial regions in all oceans and in the north Atlantic Iceland Basin region remains viable until the final months of the forecast.

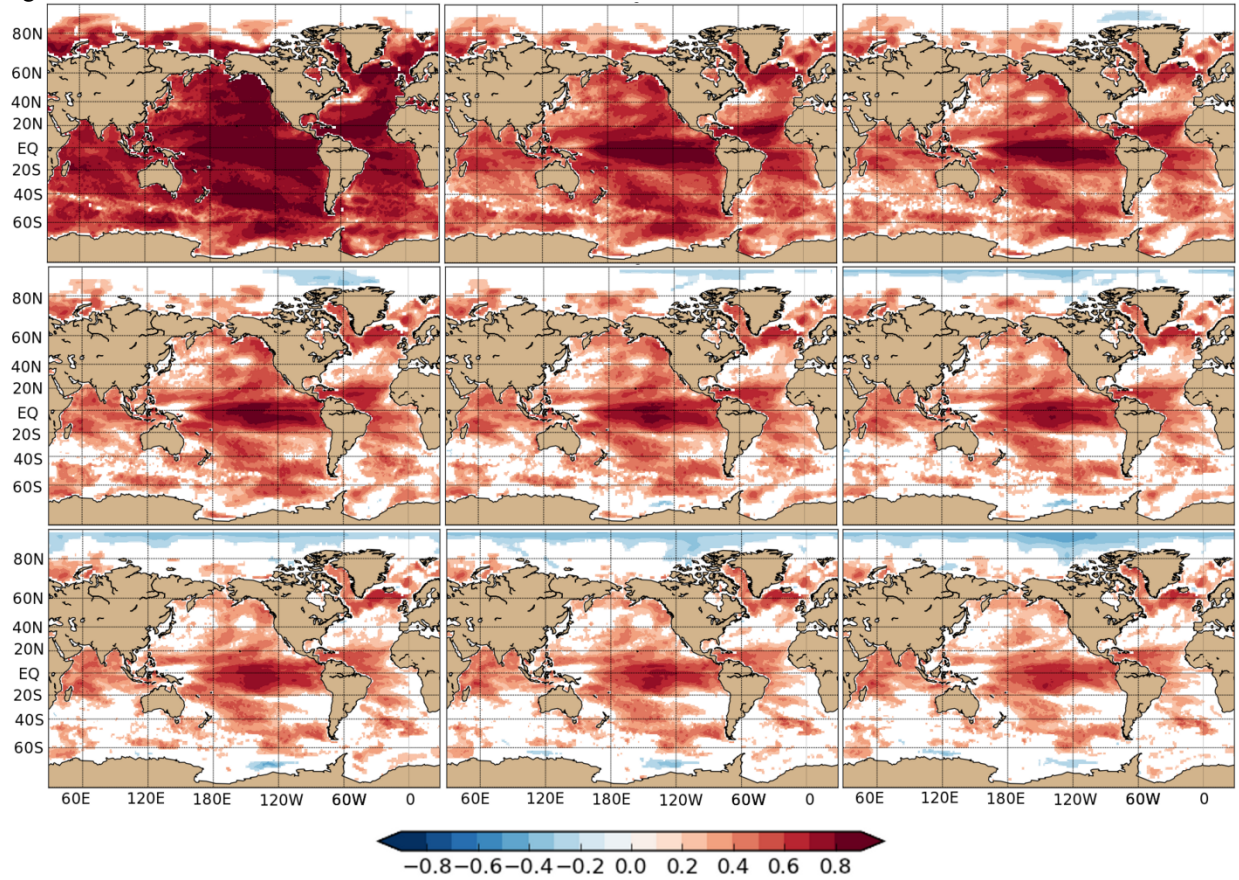


Fig 5 Global monthly SST ACC for all forecast initial months combined together; Reynolds monthly SST is used as observations; nine lead months are shown top-to-bottom, left-to-right. Pearson correlation significance test is applied with p_value at 0.01 (Pearson 1896)

4.2.2 Oceanic Indices Skill

Figure 6 and 7 show the Hovmöller diagrams of SST ACC and MSSS vs. the forecast initialization month for various SST indices. These illustrate the seasonal variability of the forecast skill. The drift period computed for the anomalies is 1993-2010 and the period used for the skill computation is 1993-2014. This period overlaps with the ensemble members that use the altimeter data. For the ACC, a Pearson correlation significance test with p -value at 0.01 is applied.

Top left panels in figure 6 show ACC of the western (Niño 4) the eastern (Niño 3) equatorial Pacific indices. Overall, albeit significant, the ACC is lower in the Niño 4 region for most of the spring and especially summer start months. In the Niño 3 region the ACC is very high (>0.8) for Jun-Sep start months, which coincides with the amplitude growth phase of El Niño/La Niña. In the Niño 4 region the highest ACC is attained during the late fall and winter forecasts, i.e. when an El Niño/La Niña is at its peak and begins to wind down. In the central Pacific Ocean (Niño 3.4 region, top row second from the right panel in Figure 6), the anomaly correlation skill is robustly high (>0.7) for February to September start months throughout the 9 month forecast. ACC drops sharply beyond May in Niño 3 and Niño 3.4 regions for most forecast started in September-January, which is an indication of the spring predictability barrier. In the Niño 4 region the spring barrier is not as pronounced, with the significant anomaly correlation skill retained through June for all forecasts initialized in late autumn and winter months (Oct-Feb). The relatively abrupt drop in ACC for forecasts starting in Jun-Aug maybe related to the rapid model drift during the first lead months for these forecasts. In the Niño 1+2 region (top right most panel in Figure 6), ACC spring predictability barrier occurs earlier than in the equatorial regions, in March, with the skill dropping below significance level after two months in January and December, and 3 months in February forecasts. The best seasons in forecasting this area are late spring and early summer.

SST forecast skill in the Indian Ocean is characterized by the presence of its own predictability barrier. This drop in the prediction skill occurs at the onset of the boreal summer monsoon and is found at both IDM poles (Waisowicz 2007). ACC skill is high beyond the first month only for the WTIO forecasts initialized in Jan-Feb and SETIO forecasts initialized in Jul-Aug. In the SETIO ACC there is a second predictability barrier is December, but the skill apparently returns later for the forecasts starting in Jul-Sep (left bottom panels in Figure 6). The IDM index defined as the difference between WTIO and SETIO indices. Forecasting the relative variability of the two regions in each of which the skill is not very robust proves to be a difficult task, although some significant skill beyond the first month is observed in the forecasts initialized in May-Nov. Here the December predictability barrier hinted at by the SETIO ACC values is strongly evident.

The TASI SST anomaly index is an indicator of the meridional surface temperature gradient in the tropical Atlantic Ocean. It was defined by Chang et al. (1997), where it was associated with a potential decadal 'dipole' mode of coupled variability in the tropical Atlantic. The GEOS-5 ability to predict the TASI values is robust but short lived: for all initialization months except Aug-Oct (and Feb-Mar), the ACC drops below the significance level after 2-3 months (bottom right panel in Figure 6). The higher skill for these forecasts may be associated with the strength of the TASI signal: the amplitude of the index is peaking during these phases of the seasonal cycle (see the bottom right panel of Fig 4).

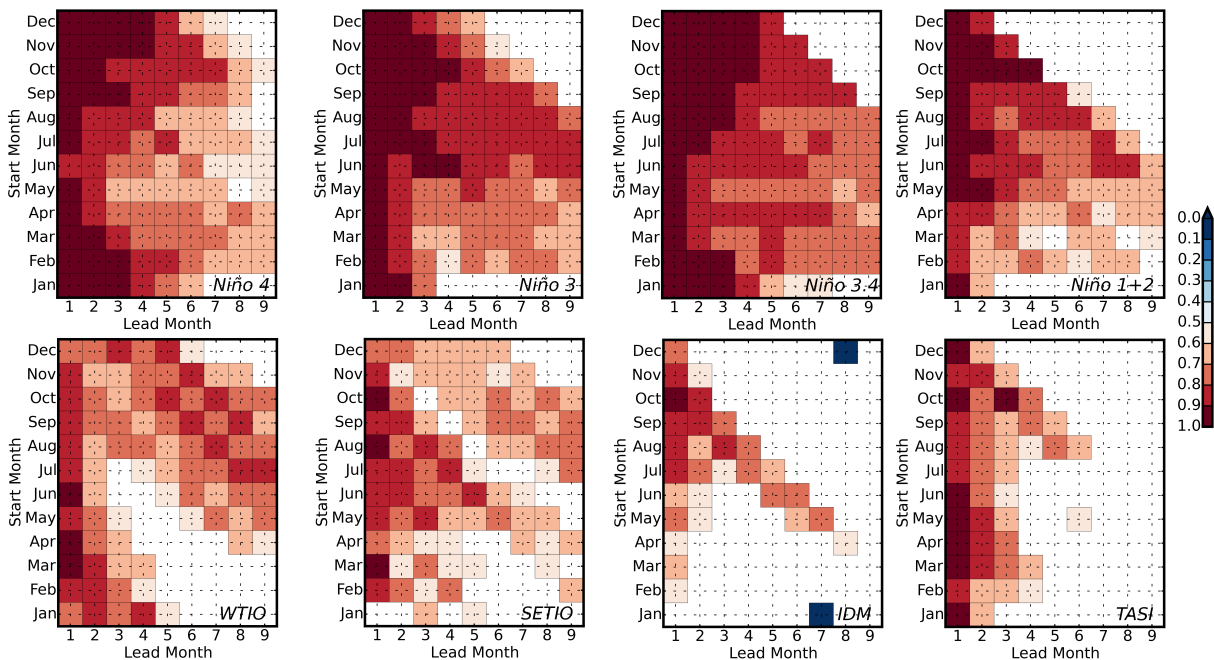


Fig 6 SST ACC for all oceanic indices. The Pearson correlation significance test with p -value at 0.01 is applied. Forecast start months are along the y-axis and lead months are along the x-axis

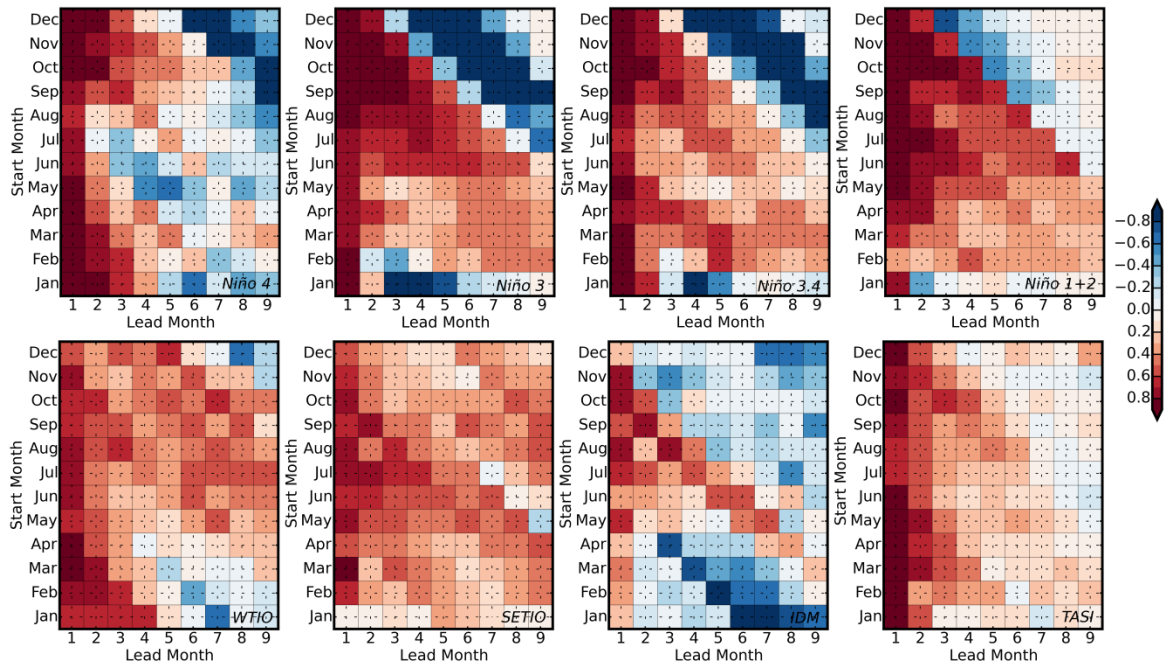


Fig 7 SST MSSS for all oceanic indices. Forecast start months are along the y-axis and forecast lead months are along the x-axis

MSSS is a characteristic of how well the anomaly amplitude is forecasted. Even when correlation skill is high, the systematic over/under-prediction of the anomaly would lead to a lower MSSS. In equatorial Pacific Ocean indices MSSS becomes negative across the spring barrier. In the Niño 4 region this can be related to an overextension of the warm pool to the west, and thus a consistent overestimation of the warm SST anomalies. Fig 1 of the Online Resource, showing the historic performance of the GEOS-5 Niño 4 index, illustrates this point: the El Niño amplitude was overestimated in 1982/83, 1991/92, 1997/98, 2002/03, 2006/07 and 2015/16 cases. Additionally, and this is evident in all three equatorial indices, the system tends to falsely predict a warming trend (as opposed to the neutral condition in reality) for the following spring/summer for forecasts starting in boreal winter. This contributes to the drop in anomaly correlation skill (spring barrier) and the low amplitude skill.

For the Indian Ocean, while ACC/MSSS skills for the Western and Southeastern indices appear to be significant/positive for most of the forecasts, both skills for the IDM index are low except for the May-November starts, and even for these, the predictable lead time is 2-5 months. This is comparable to other dynamic models (Shi et al 2012). The Tropical Atlantic Ocean index skill shows forecast outperforming climatology in terms of error absolute value, as well as anomaly correlation, for the short term predictions (2-4 months).

4.2.3 Case Study: Major El Niño Event of 15/16

Figure 8 shows spaghetti plots of the ensemble mean forecasts for each start month for 2015-2016 (in color). Observations from Reynolds SST are shown by a solid black line, ocean analysis is shown by a dashed black line. The color scale represents the ratio between the forecast absolute departure from the observations and the standard deviation of the ensemble at that particular lead time. High values of this measure may be indicative of ensemble underdispersion. The 2015/2016 El Niño was considered a Central Pacific event so of all the indices, the Niño 4 index in the Western Pacific exhibited the smallest observed anomaly compared to other Pacific Ocean indices and Niño 3.4 had the highest observed anomaly. GEOS-5 overpredicted the magnitude of the SST anomaly at the peak of the El Niño in the western central Pacific (Niño 4 index) by as much as 1.5°C. The timing was also missed by summer and fall (Jun-Nov, 2015) forecasts, they all showed the maximum in January 2016, while it occurred in Nov 2015. So great was the forecasts departure from the observations, that the latter barely fit within the ensemble envelope in October, 2015 through January, 2016, the ensemble mean being as far from the observations as 4 standard deviations of the ensemble. The maximum of the cooling phase was also overpredicted by more than 1°C,

and the timing was too early: the winter and spring (Jan-Jun, 2016) forecasts showed the lowest temperature in August 2016, while in reality, the cooling gradually took place over the course of 2016. GEOS-5 accurately predicted surface warming in the Niño 3 region as early as March 2015. The following forecasts, starting in boreal summer (Jun-Aug, 2015), showed the warming being too early by about 2 months, however the timing of the peak SST anomaly in November 2015 - January 2016 was predicted well by all spring and summer forecasts except August 2015, which showed the peak in February 2016. This was also the warmest of all the predictions; the rest of the forecasts were within 0.5°C of observations, which corresponds to roughly one standard deviation of the ensemble. The amplitude of the cooling following the El Niño peak in this region was overpredicted by the forecasts initialized in April and May 2016 (they called for a moderate La Niña), while the rest of the forecasts, earlier (Jan-Mar, 2016) and later (June, 2016 onwards) predicted neutral conditions.

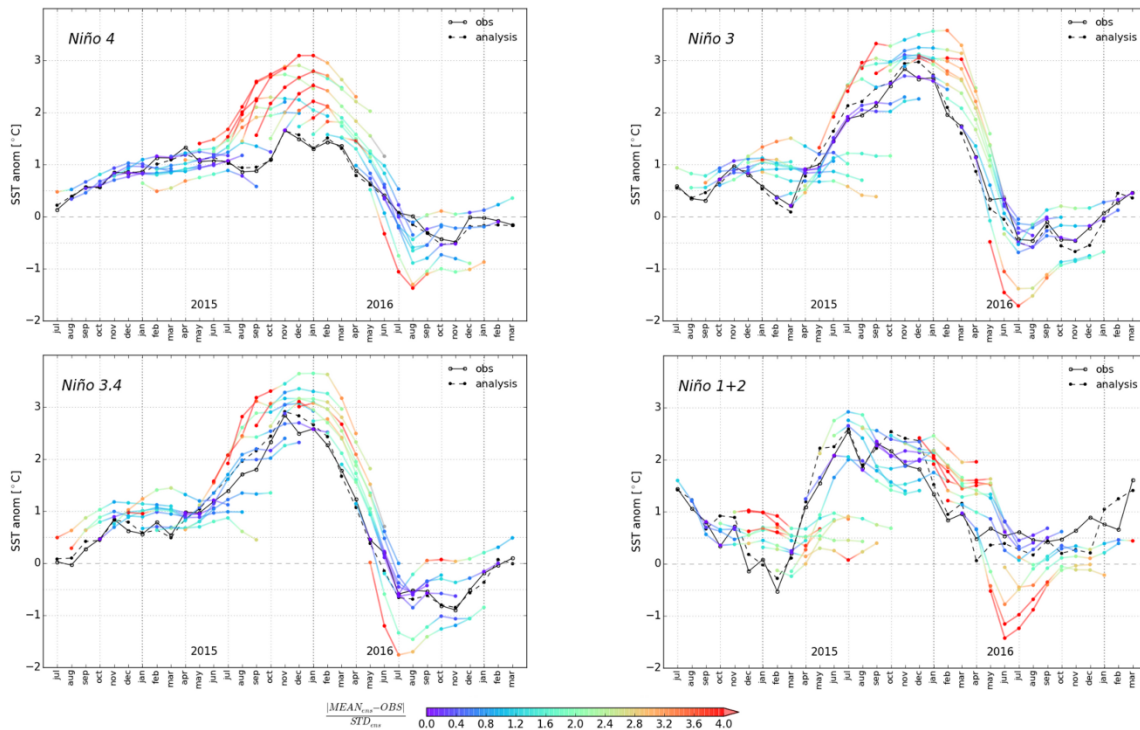


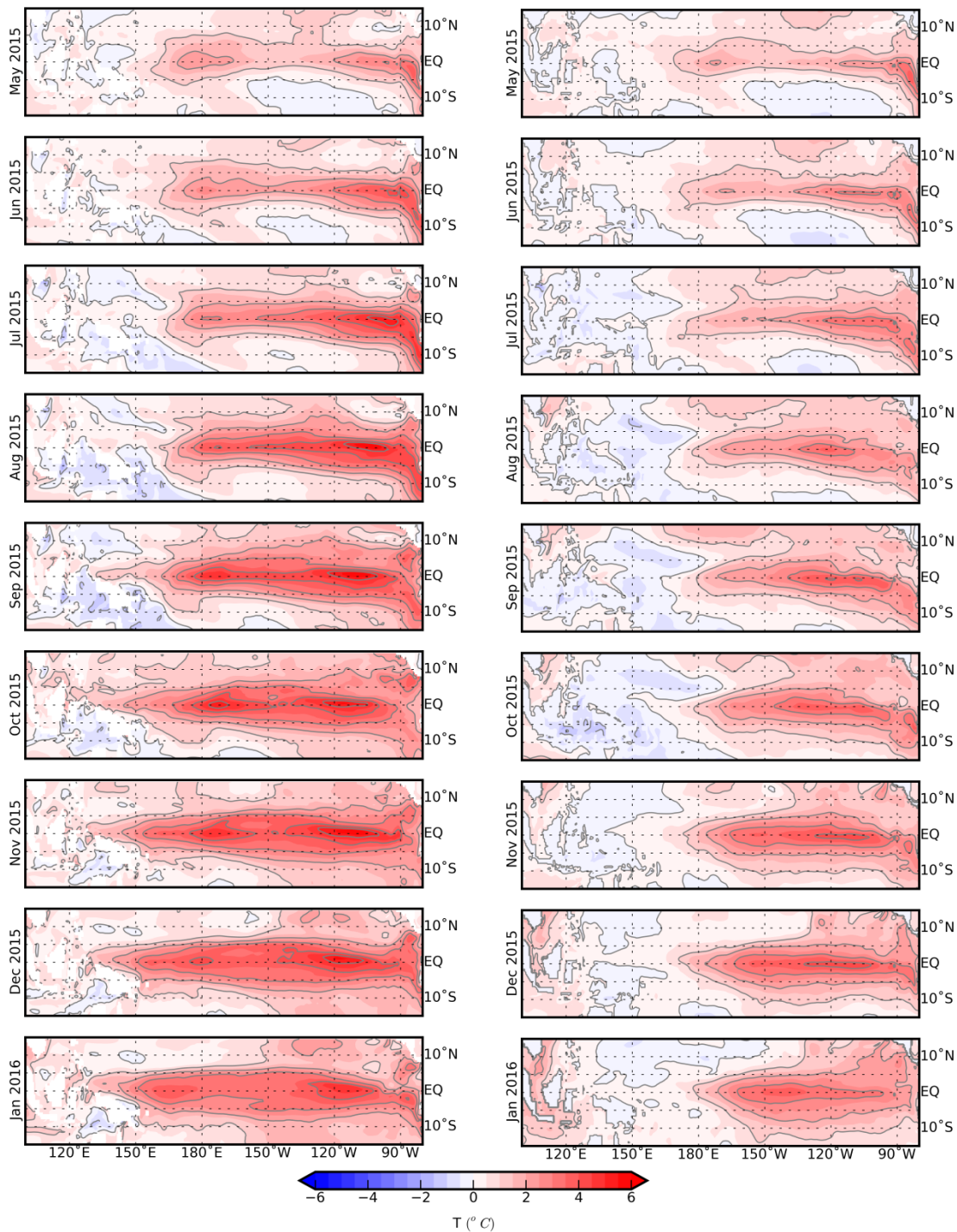
Fig 8 Niño 4, Niño 3, Niño 3.4 and Niño 1+2 monthly mean SST forecasts; the solid color lines show the ensemble mean, the black line is observations (Reynolds SST), the dashed black line is the ocean analysis from which the initial conditions for the forecasts were generated. The color scale represents the ratio between the forecast absolute departure from the observations and the standard deviation of the ensemble at that particular lead time

Note the peak of the 2015/2016 El Niño event that occurred in the NDJ season for the equatorial indices. The GEOS-5 model predicted the correct timing and magnitude of this peak for the Niño 3.4 index starting in February 2015. This index is the one most widely used for ENSO forecasting, thus intercomparison between various models is readily available. GEOS-5 model performs similarly to most other models involved in NMME, although more often than not, it tends to have stronger ENSO events than other models.

The Niño 1+2 region reached its peak anomaly in July 2015. The GMAO model predicted the timing of this event starting in March but underestimated the magnitude by 0.5°C. The June forecast was very close to the observations in timing and magnitude. In all regions, the forecast of the cooling phase of the ENSO starting in May, 2016 was exaggerated for all four indices. By July, 2016, start time, the equatorial indices forecasts picked up the transition to the neutral conditions.

Fig 9 shows the evolution of the SST and the next figure (Fig 10) shows the equatorial subsurface temperature during the onset of the 2015/2016 El Niño. The overextension of the warm water anomaly to the west of the date line clearly shows the forecast difficulty in the Niño 4 region. It is consistent with lower skills in this area, as noted in Ham et al (2014b). The GEOS-5 system exhibited similar behavior during the previous ENSO events (1997/98 El

437 Niño and 1982/83 El Niño, see Online Resource 1-6 illustrating the historic SST indices values in the GEOS-5
 438 forecasts).



439 **Fig 9** Evolution of the equatorial Pacific Ocean SST during the onset of the 2015-16 El Niño. Left panel is the monthly mean
 440 forecast SST from May 2015 initial conditions. Right panel is the concurrent MERRA ocean analysis
 441

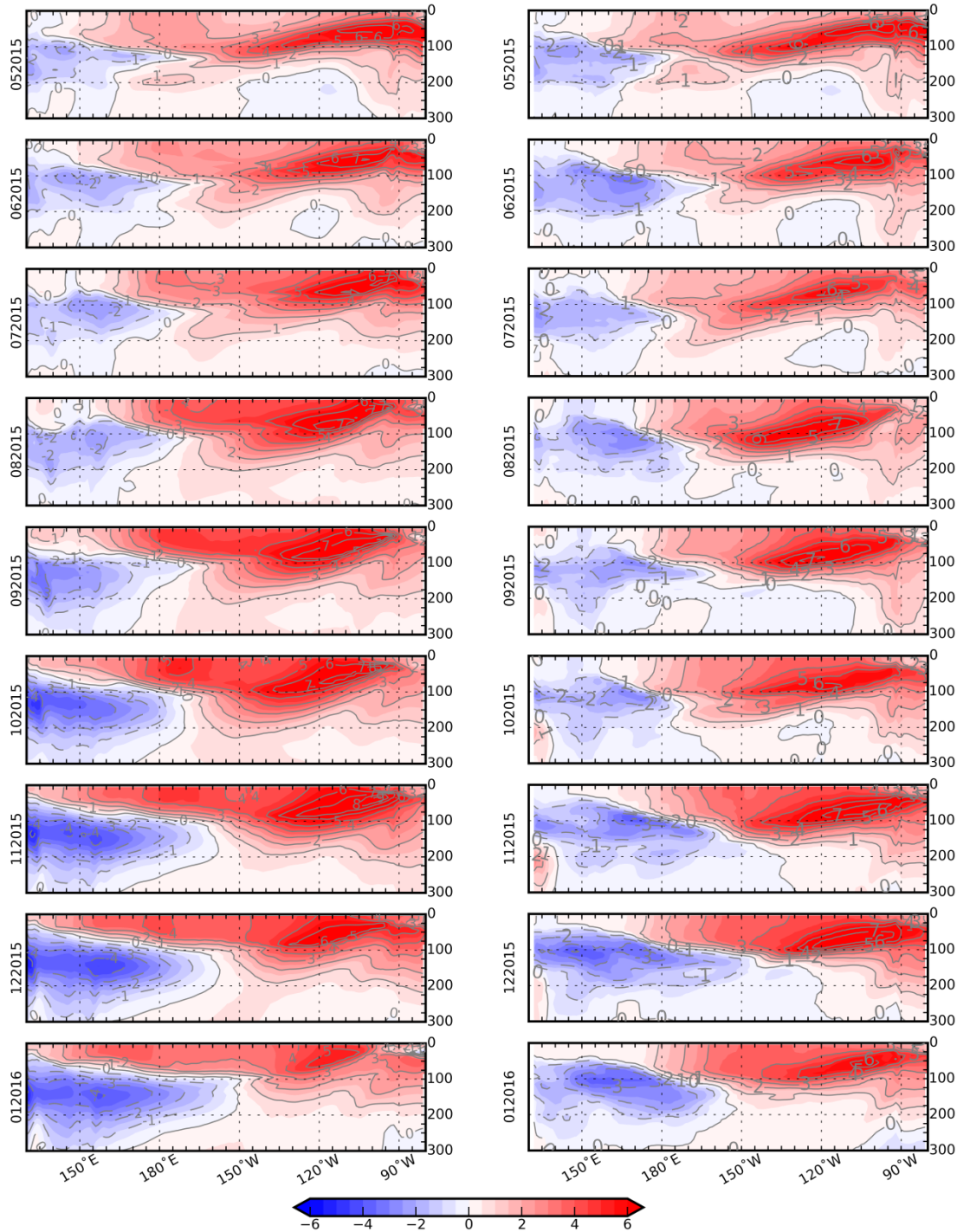


Fig 10 Evolution of the subsurface equatorial Pacific Ocean temperature during the onset of the 2015-16 El Niño. Left panel is the forecast monthly mean T from May 2015 initial conditions. Right panel is the concurrent MERRA-ocean analysis

5 Forecast Bias and Skill: T2M and Precipitation

Similarly to the SST, but for temperature at 2 meters (T2M) and precipitation, we first present the global forecast bias maps and then discuss the regional skills and case studies of two extreme events.

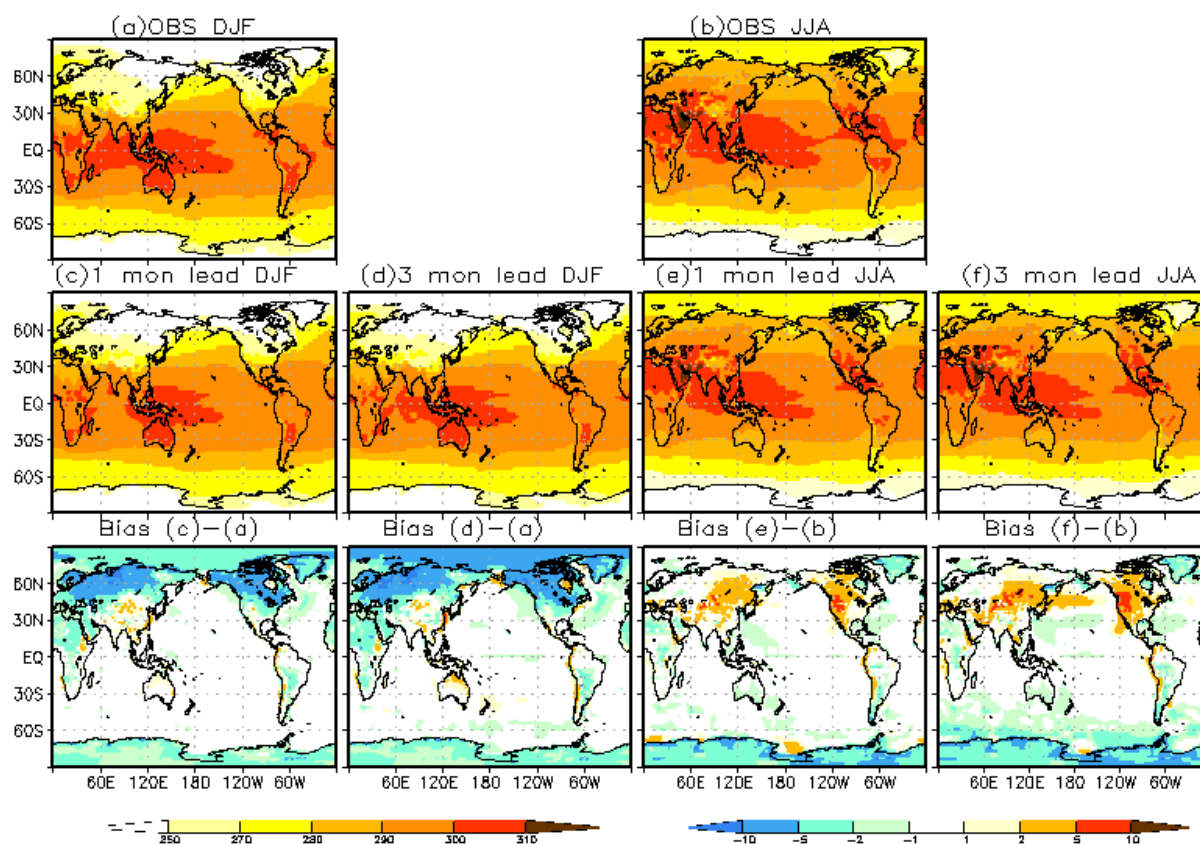


Fig 11 An example of 2m air temperature seasonal forecast bias for 1 and 3 months lead times. Winter and summer observed and predicted fields are shown in the two top rows. The bottom row shows the differences between the model and the observations

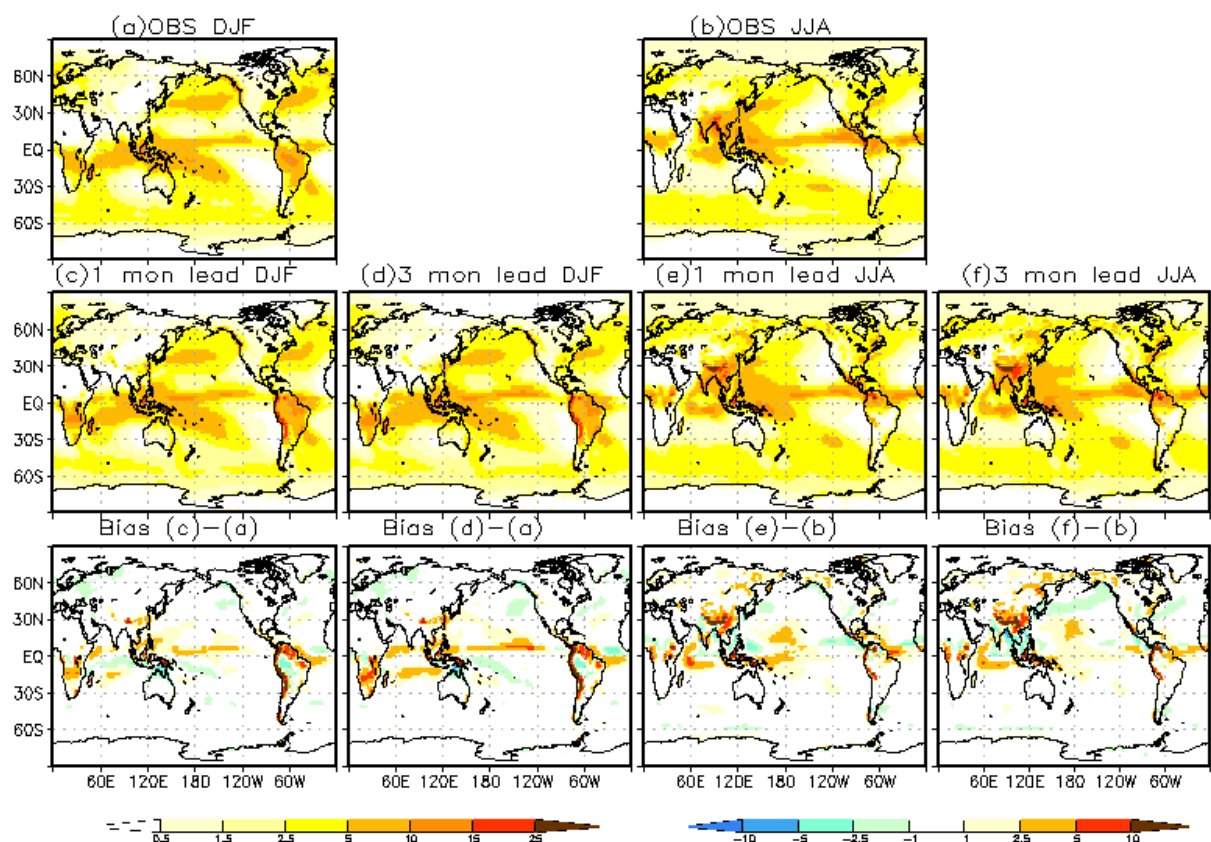


Fig 12 An example of precipitation seasonal forecast bias for 1 and 3 months lead times. Winter and summer observed and predicted field are shown in the two top rows. The bottom row shows the differences between the model and the observations

The bias shown in Figures 11-12 is the systematic departure in predicted and observed climatology during the 30 year (1982-2011) period. For both T2M and precipitation, MERRA-2 (Bosilovich 2016, Molod 2015) data was used as the observational validation reference. The first lead month and the third lead month bias for winter and summer are showed. The largest bias for T2M is over the winter Arctic Ocean and it increases with lead time. GEOS-5 underestimates the sea ice temperature in winter. For the northern hemisphere summer, GEOS-5 tends to overestimate T2M over land, especially in Asia and the west coast of North America. Overall the bias for T2M remains small over the tropical oceans in all seasons over the full length of the forecast. The bias for precipitation varies slightly with lead time for both winter and summer. Larger bias is found in the summer than in the winter over land in the northern hemisphere. This is possibly due to the fact that summer precipitation over land is more likely to be affected by regional and local factors, thus uncertainty in model parameterization as vegetation cover, cloud physics etc. could play larger roles in the precipitation bias.

5.2 Forecast Skill

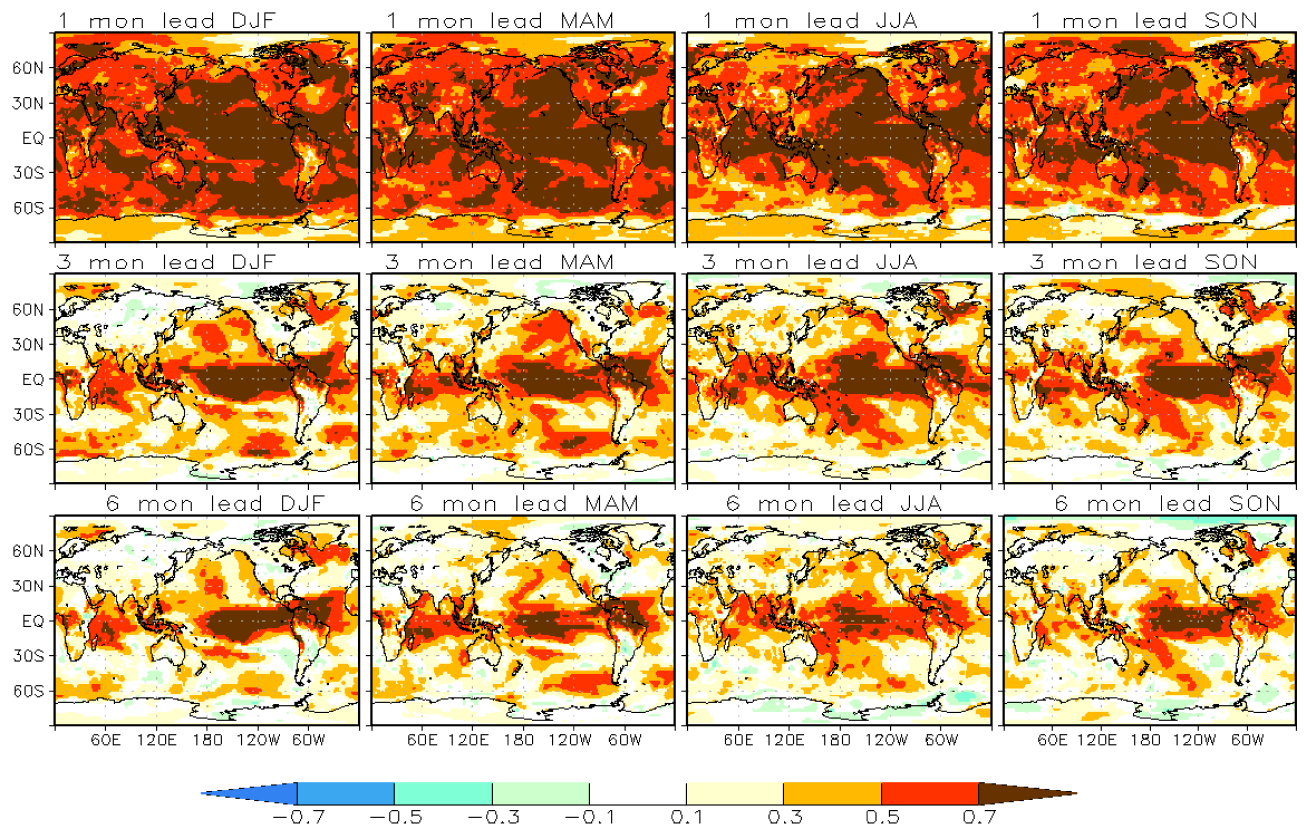


Fig 13 Monthly mean 2m air temperature anomaly correlation for seasonal forecast and observations (MERRA-2)

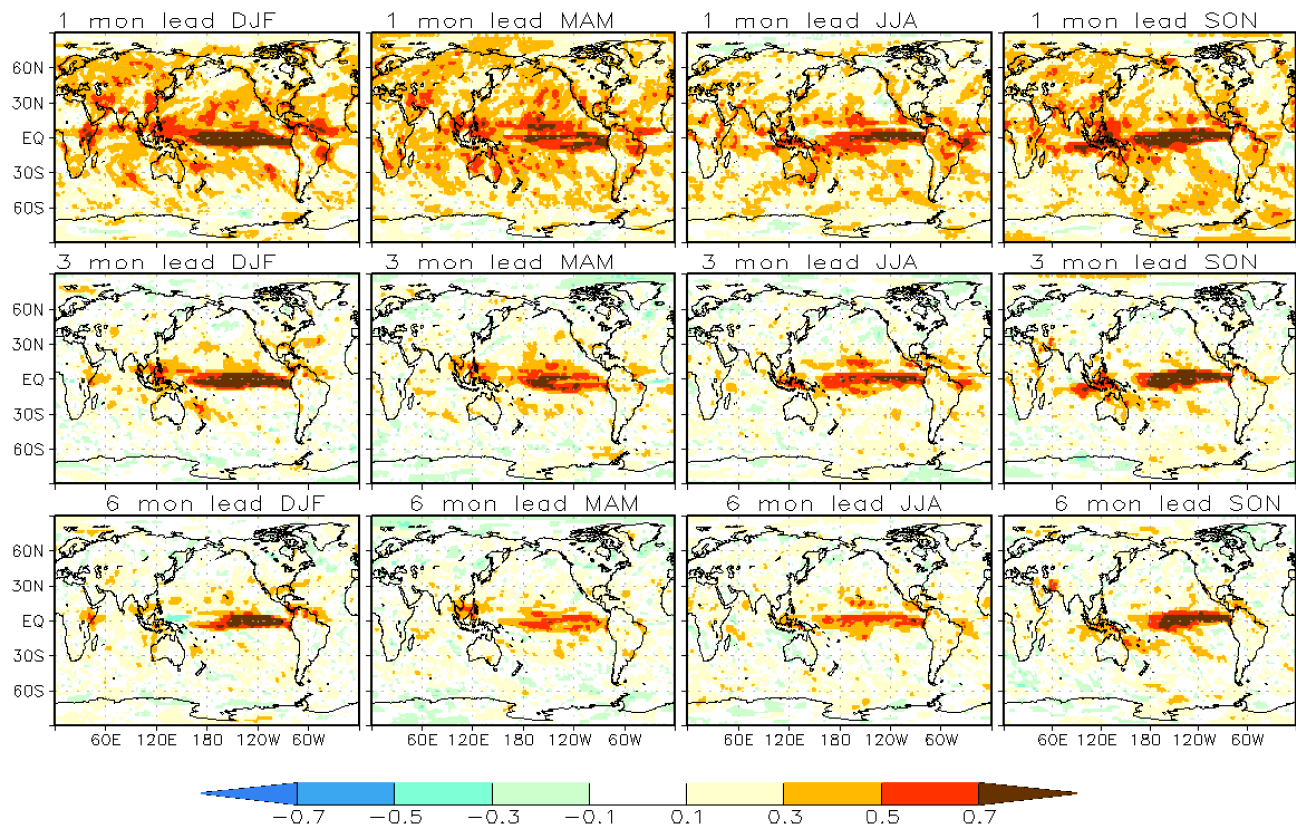


Fig 14 Monthly mean precipitation anomaly correlation for seasonal forecast and observations (MERRA-2)

Seasonal skills of T2M (Fig 13) and precipitation (Fig 14) are calculated as anomaly correlations between GEOS-5 forecast and observations. GEOS-5 performs well for T2M for the first lead month for all seasons, especially over the tropical oceans. Although the T2M skills over the tropical oceans remains high even after six months, the skills over land diminish quickly after the first lead month. Precipitation skills are generally lower than those of T2M, and there is hardly any skill after the first month for the extra tropics. However, over the East Pacific Ocean, where ENSO has a dominant influence on precipitation, the anomaly correlation remains high until the sixth lead month.

5.2.1 Regional Average Skills and Case Studies

For the discussion in this section we consider special regions of particular socio-economic interest: the Amazon basin bounded by 80°W-50°W, 20°S-10°N, the Great Plain between 30°N-50°N, 110°W-100°W and Southern India between 10°N-15°N, 77°E-80°E. These areas are shown by pink rectangles in Fig 1 and labeled Amazon, GP and SI respectively.

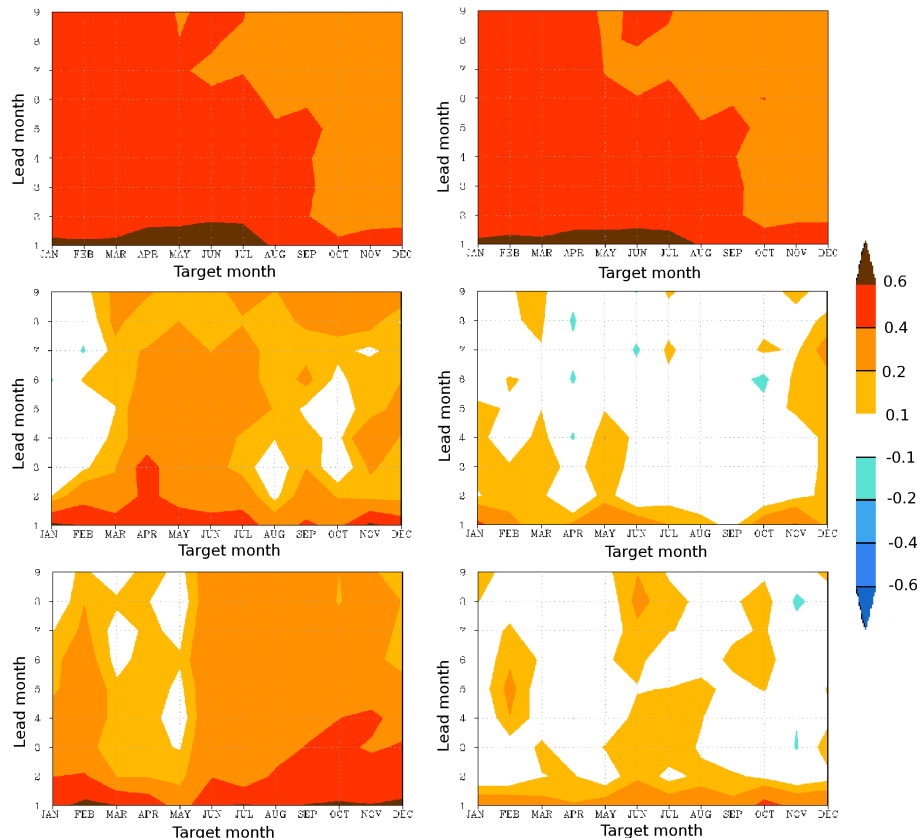


Fig 15The anomaly correlation for T2M (left) and precipitation (right) between the forecast and MERRA-2 for Amazon River basin (top), Great Plain (middle) and Southern India (bottom) regions. Target month (x-axis) represents the date of the forecast, and the lead month (y-axis) represents how long that forecast was in months, i.e. target month May with lead 4 means May forecast initialized in February

A closer look (Fig 15, top row) at the broad region encompassing the Amazon River basin reveals a good overall skill in terms of anomaly correlation for both T2M and precipitation throughout the seasonal forecast. For the initialization months Jan-May, it stays above 0.4 up to lead month 9 and for Jun-Jul it starts above 0.6 for the first month and drops below 0.4 only after month 6.

Seasonal forecasts serve as an important prediction tool for extreme events. At the current stage, there are uncertainties and difficulties in seasonal forecasts to accurately capture certain events. However, it is intriguing to see how GOES-5 performs in various extreme events. Two examples of such extreme events will be discussed next: drought over the Great Plains in 2012 and flood over the southern India coast in 2015. Figure 15 (middle and bottom panels) shows the T2M and precipitation anomaly correlation skill in these regions. One can see that there is little correlation between the observations and the predicted precipitation, yet the strong signal during the extreme events may give an opportunity for the forecast to capture its characteristics.

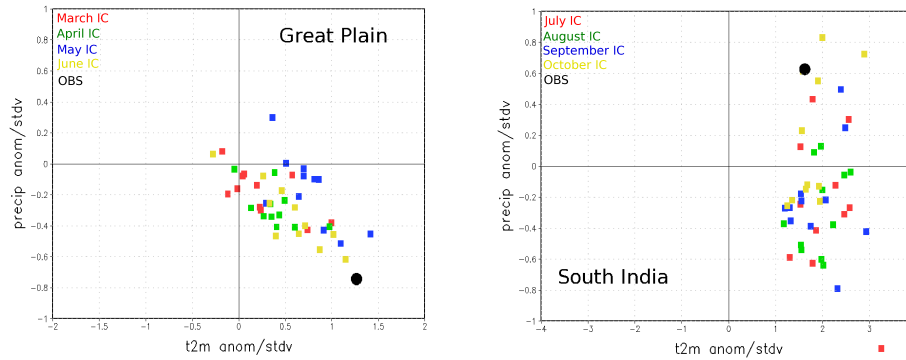


Fig 16 Case studies of the drought over the Great Plains in 2012 and the flood over the Southern India coast in 2015. One to four month lead forecasts of T2M and precipitation are plotted against the observations (MERRA-2). For the Great Plain the target month is July 2012, for the South India region the target month is November 2015, The anomalies are standardized using the corresponding standard deviation

In the scatter plots in Fig16, one to four month lead forecasts of T2M and precipitation are plotted against the observations for the two events. The anomalies are standardized using the corresponding standard deviation. In the case of the Great Plain drought, GEOS-5 underestimates the deficit in both temperature and precipitation for all lead months. The underestimations are most obvious in the spring initialized forecasts and gradually decrease closer to summer. By May and June, GEOS-5 clearly predicts hotter and drier conditions for that summer, although the magnitude of the drought is less than in observations. The Great Plain is a region where the local water cycle is sensitive to the land surface representations and therefore is sensitive to land initializations. This characteristic of the region is also presented in this scatter plot. For the case of the southern India flood, GEOS-5 shows a much larger model spread for every initialization month. However, the large model spread highlights the low predictability of the seasonal forecast. The southern India flood is highly related to the location and strength of the Indian winter monsoon. The winter monsoon predictability therefore highly restrains the performance of the GEOS-5 seasonal forecast for floods.

6 Sea Ice Outlook

Seasonal forecasting systems focus on the ability of coupled atmosphere/ocean models to predict variability in the tropical Pacific Ocean and its associated higher latitude teleconnections (Kirtman et al. 2014). For middle and high latitudes, predictability derived from local oceanic sources has been thought to be limited (e.g., Barsugli and Battisti 1998). But Arctic sea-ice cover has a decorrelation time scale of up to 5 months (Blanchard-Wrigglesworth et al. 2011). Moreover, model experiments have indicated sea-ice predictability on seasonal time scales and longer, with indications of signal re-emergence beyond one year (Holland et al. 2010; Tietsche et al. 2014; Guemas et al. 2014). The prospect of a predictability reservoir has received considerable interest (Richter-Menge et al. 2012; Stroeve et al. 2014; Hamilton and Stroeve 2016). Potentially, seasonal forecasts of sea ice have utility for a variety of human endeavors including commerce, mineral exploration, and indigenous activities (Stroeve et al. 2015). Arctic sea-ice extent is also considered a climate variable. Mechanisms controlling its variability and trend are the subject of extensive observational and modeling studies (e.g., Perovich and Richter-Menge 2009; Vaughan et al. 2013). The presence of floating ice on the ocean radically alters surface properties; it has immediate influence on the exchange of energy and moisture between the ocean and the overlying atmosphere. Model experiments have demonstrated the impact of ice cover on regional Arctic climate, including air temperature and precipitation (Deser et al. 2010; Alexander et al. 2004), and studies have also suggested an influence on large-scale conditions extending beyond the immediate Arctic Basin (Thomas et al. 2014). In 2008, a challenge was formulated for comparing and evaluating experimental seasonal predictions of the September Arctic sea-ice extent (ARCUS, 2008), which became known as the Sea Ice Outlook. Many of the seasonal forecasting systems involved with NMME have participated, including the GMAO. Results have been mixed. Assessments have shown that predictions have reduced skill in years where the observed ice cover departs significantly from the long-term trend (Hamilton and Stroeve 2016). Subsequent evaluation of models participating in the Sea Ice Outlook has shown that forecasts have difficulty surpassing the skill of damped persistence, and difficulty predicting each other (Blanchard-Wrigglesworth et al. 2015).

Forecasts submitted to the Sea Ice Outlook were composed of the ensemble members initialized at every five days of the prior month, along with ten members initialized at the beginning of the month. Unlike the ENSO forecasts and the ensemble members submitted to NMME, the sea ice forecasts made use of three experimental members, which were produced in hindcast mode beginning in 1993. These experimental members featured the inclusion of altimeter data, which was obtained from the Archiving, Validation and Interpretation of Satellite Oceanographic data project (AVISO; <http://www.aviso.altimetry.fr/>) and used in the ocean assimilation.

SIPN/ARCUS June Sea Ice Outlooks of September Mean Extent: 2014-2016

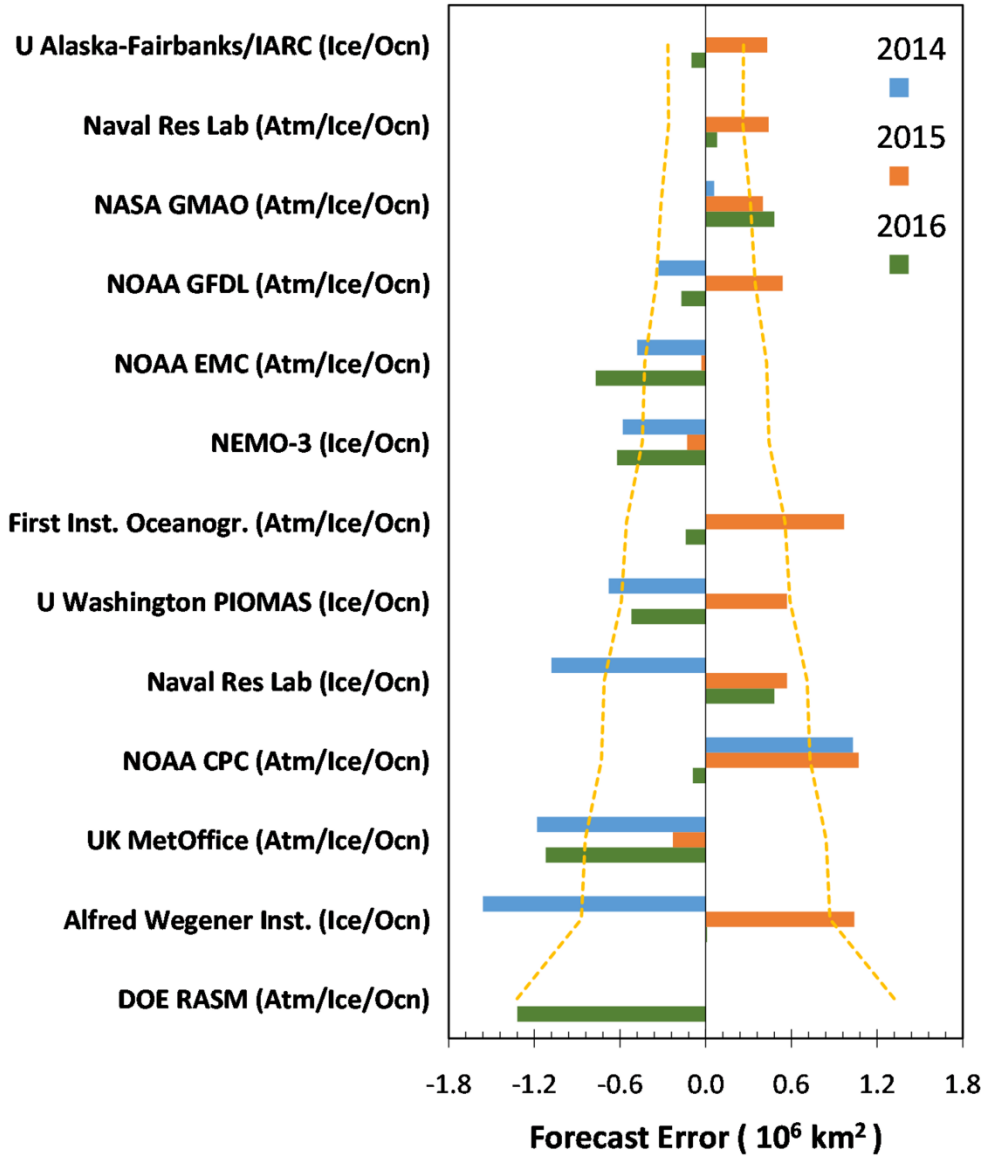


Fig 17 Comparison of forecast September Arctic ice extent error from submitted June Sea Ice Outlook models for the period 2014-2016. Yellow dashed line indicates the average error over the three years. Error is computed as the difference of the forecast value minus the extent from passive microwave data (Cavalieri et al. 1996)

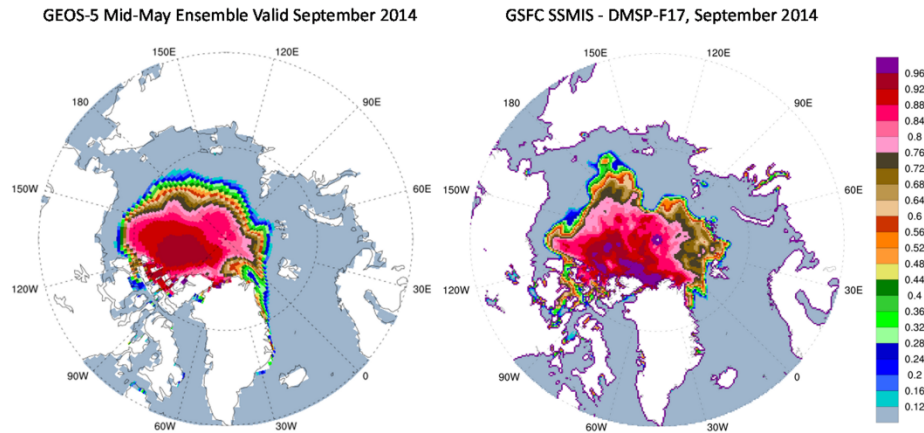


Fig 18September 2014 forecast and the observed spatial pattern of Arctic sea ice

As shown in Fig 17, the GMAO system has performed well in comparison to other models over the period in which it has participated in the Sea Ice Outlook. For the previous three Outlooks, the average extent error is $0.32 \pm 0.22 \times 10^6 \text{ km}^2$ for the GMAO system as compared to $0.57 \pm 0.42 \times 10^6 \text{ km}^2$ for the average of all dynamical predictions. The uncertainty denotes the standard deviation of the forecast errors. Fig 18 also indicates that the spatial patterns for the September 2014 forecast were similar to the observed pattern. Over the hindcast period of 1998 to 2015, the June forecast explains 49 percent of the observed September ice extent variance, which may be considered of marginal skill. But this belies several critical issues with the forecast system, which are largely associated with initial conditions. As previously noted, the MERRA atmospheric reanalysis is used in the ocean assimilation. Cullather and Bosilovich (2012) found near-surface air temperatures in MERRA are as much as 10°C too warm in the late Arctic spring, owing to an erroneously low, fixed sea-ice albedo used in the uncoupled atmospheric reanalysis. The surface temperature bias and its effects on the GEOS-iODAS oceanic temperatures led to an anomalous reduction in forecast ice cover initialized during early summer months. Fig 20 indicates the increase in forecast error in summer months, such that forecast skill actually becomes reduced with decreasing lead time. This inhibits contributions to the Sea Ice Outlook for one- and two-month lead times to the September Arctic ice minimum. The June Outlook contributions are based on the prior month's initialization. Analysis of hindcasts also finds low ice cover for spring forecasts initialized during the first ODAS data stream covering the period until 1993, as shown in Fig 19. Low ice volume associated with this stream and the interaction with the erroneously warm atmospheric forcing in the analysis results in poor ice forecasts over the time period. Hindcast skill improves in the later GEOS-iODAS stream and with the introduction of altimetry-based forecast ensemble members after 1993; as seen in Fig 19, variability in the ensemble mean forecast is comparable to observed values after 1998. The period from 1998 to the present is used for a simple bias correction in the forecast to account for differences with the NSIDC Sea Ice Outlook. This is a common practice among Sea Ice Outlook participants. Over the full period, forecast ice cover for the Southern Ocean is patchy and not comparable to observation.

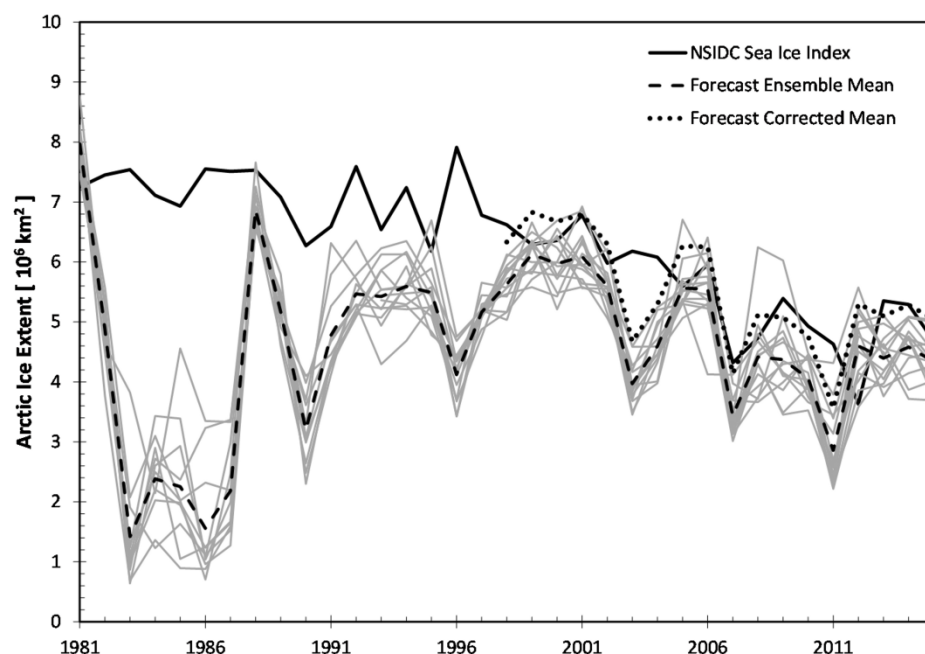


Fig 19 September mean sea ice extent from NSIDC Sea Ice Index (solid black line; Fetterer et al. 2016), and from ensemble members of the June forecast (grey lines). The ensemble mean is indicated with a black dashed line, and a bias-correction is indicated with a dotted line, in 10^6 km^2

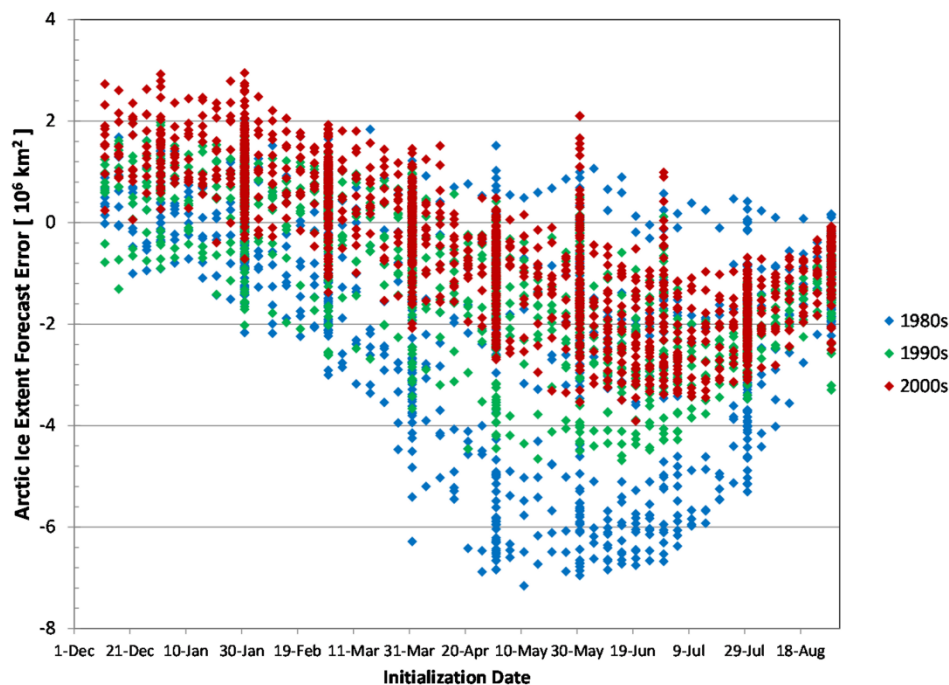


Fig 20 September sea ice extent from each ensemble member minus the NSIDC Sea Ice Index observed value, plotted based on the initialization date. Colors indicate the year of the forecast

The lessons learned from this initial seasonal forecasting system exercise are useful in the construction of new forecast systems. First, the contrast in the competitiveness of the system shown in Figures 17-18 with the difficulties indicated in Figures 19-20 suggest the continued experimental nature of sea-ice forecasts on these time scales, but

also that some significant improvement is relatively straightforward – for example, with improved atmospheric temperature forcing of the ocean analysis in the melt season such as in MERRA-2 (Bosilovich et al. 2016). The utility of hindcasts for the general characterization of the sea-ice forecast system suggests that the anomaly forecasting approach of NMME is also advantageous for the Arctic. But the simple bias correction approach that has been heretofore used is likely problematic. Bias correction of the hemisphere-summed ice extent is a means of addressing an inadequate representation of seasonality in the model, but it is mostly used here to address the mismatch of the land-sea mask between various models and observing system grids (see for example, Blanchard-Wrigglesworth et al. 2016). A better methodology for sea-ice intercomparison between models and observations is required. The issues shown here emphasize the role of initial climate conditions for improved forecasts. This includes an investigation for improving the representation of sea-ice characteristics in the analysis state.

7 Conclusions

In this study we provide the details of the GEOS-5 seasonal forecast system setup, which is used in particular to provide monthly contributions to the NMME project. The SST, T2M, precipitation and sea ice extent skills are documented for the comparison with other systems and to track the differences, hopefully, improvements, with the future system currently under development. Notable problems in the current system are the large SST drift in the northern Atlantic Ocean (strongest for the forecasts initialized in winter and spring), in the equatorial Pacific Ocean (strongest for the forecasts initialized in winter and spring), in the northern Pacific Ocean (for the forecasts initialized in late fall and winter) and in the Southern Ocean (for the forecasts initialized in austral winter). Anomaly correlation SST skill is poor in the western Pacific Ocean (Niño 4 index) relative to the central and eastern equatorial regions (Niño 3 and Niño 3.4 indices). This is related to the overextension of the warm water anomaly to the west during the El Niño events. Ham et al (2014b) attributes this to the weak thermal damping and the erroneous zonal advective feedback as a response of the wind-driven current to the wind forcing. The strong ocean sensitivity to the wind forcing may be mitigated in the future by coupling the ocean with the new improved atmospheric analysis (MERRA-2), in which the surface winds are significantly improved compared to MERRA (Bosilovich et al, 2016), in particular, over the western equatorial Pacific Ocean.

The largest (negative) bias for T2M is over the winter Arctic Ocean related to the fact that GEOS-5 underestimates the sea ice temperature in winter. The largest positive T2M bias occurs during the northern hemisphere summer over land, especially in Asia and the west coast of North America. T2M is not significantly biased over the tropical oceans. The largest precipitation bias is found in the summer over land in the northern hemisphere, likely related to the uncertainty in model parameterization of vegetation cover, cloud physics etc. Anomaly correlation T2M skill is high for at least 6 lead months for all seasons over the tropical oceans, but drops quickly (after 1 month) over land. Significant precipitation skills in terms of anomaly correlation are found only over the equatorial eastern Pacific Ocean, where ENSO has a dominant influence on precipitation. Everywhere else the anomaly correlation drops to near zero after the first month of the forecast.

The experimental sea ice forecast provides a benchmark for the future system evaluation. The current GEOS-5 system is on par with other comparable models based on the 3 year comparison within the Sea Ice outlook project. Yet there are known shortcomings in the sea ice initialization and ocean and atmospheric feedbacks. In addition to providing better forcing to the ocean model via MERRA-2 analysis, the sea ice forecast has a potential to benefit from assimilating new types of observations during the ocean and sea ice initialization procedure, such as sea surface height and ice thickness.

The GEOS-5 seasonal forecast system has been in service since early 2012. Since its inception, new versions of the models have become available, and the new ensemble ocean and sea ice assimilation system that is capable of processing new data types has been put in place. As the next system is being developed, this paper will be among those providing reference for evaluating its performance.

Acknowledgments

CMAP precipitation fields were obtained from the NOAA Climate Prediction Center. Ocean altimetry data were obtained from AVISO. Reynolds SST fields are obtained from the NOAA National Climate Data Center. Passive microwave sea ice concentration was obtained from the National Snow and Ice Data Center. Sea Ice Outlook forecast values were obtained from the Arctic Research Consortium of the US. This work is funded through the NASA Modeling, Analysis and Prediction (MAP) program.

636

637 **References**

- 638 Alexander MA, Bhatt US, Walsh JE, Timlin MS, Miller JS, Scott JD (2004) The atmospheric response to realistic
639 Arctic sea ice anomalies in an AGCM during winter. *J Climate* 17:890-905 doi:10.1175/1520-
640 0442(2004)017<0890:TARTRA>2.0.CO;2
- 641 Antonov JI and coauthors (2010) *World Ocean Atlas 2009, Volume 2: Salinity*. Levitus S, Ed. NOAA Atlas
642 NESDIS 69, U.S. Government Printing Office, Washington, D.C., 184 pp
- 643 ARCUS (2008) *Arctic Observation Integration Workshops Report, 17-20 March 2008, Palisades, New York, USA*.
644 SEARCH Project Office, Arctic Research Consortium of the United States (ARCUS), Fairbanks, Alaska, 63 pp
- 645 Bacmeister J, Suarez M, Robertson F (2006) Rain Reevaporation, Boundary Layer–Convection Interactions, and
646 Pacific Rainfall Patterns in an AGCM. *J AtmosSci* 63:3383–3403, doi: 10.1175/JAS3791.1
- 647 Bailey D, Holland M, Hunke E, Lipscomb B, Briegleb B, Bitz C, Schramm J (2010) *Community Ice Code (CICE)*
648 *User's Guide Version 4.0*. National Center for Atmospheric Research, Boulder, Colorado, 22 pp
- 649 Balmaseda MA, Hernandez F, Storto A, Palmer MD, Alves O, Shi L, Smith GC, Toyoda T, Valdivieso M, Barnier
650 B, Behringer D, Boyer T, Chang YS, Chepurin GA, Ferry N, Forget G, Fujii Y, Good S, Guinehut S, Haines K,
651 Ishikawa Y, Keeley S, Köhl A, Lee T, Martin MJ, Masina S, Masuda S, Meyssignac B, Mogensen K, Parent L,
652 Peterson KA, Tang YM, Yin Y, Vernieres G, Wang X, Waters J, Wedd R, Wang O, Xue Y, Chevallier M, Lemieux
653 JF, Dupont F, Kuragano T, Kamachi M, Awaji T, Caltabiano A, Wilmer-Becker K, Gaillard F (2015) The Ocean
654 Reanalyses Intercomparison Project (ORA-IP). *Journal of Operational Oceanography*. 8 Iss. sup1
- 655 Barsugli J, Battisti D (1998) The basic effects of atmosphere-ocean thermal coupling on midlatitude variability. *J*
656 *AtmosSci* 55:477-493 doi:10.1175/1520-0469(1998)055<0477:TBEAO>2.0.CO;2
- 657 Blanchard-Wrigglesworth E, Armour K, Bitz C, DeWeaver E (2011) Persistence and inherent predictability of
658 Arctic sea ice in a GCM ensemble and observations. *J Climate* 24:231-250 doi:10.1175/2010JCLI3775.1
- 659 Blanchard-Wrigglesworth E, Cullather RI, Wang W, Zhang J, Bitz CM (2015) Model forecast skill and sensitivity to
660 initial conditions in the seasonal Sea Ice Outlook. *Geophys Res Lett* 42:8042-8048 doi:10.1002/2015GL065860
- 661 Blanchard-Wrigglesworth E, Barthélemy A, Chevallier M, Cullather R, Fučkar N, Massonnet F, Posey P, Wang W,
662 Zhang J, Ardilouze C, Bitz CM, Vernieres G, Wallcraft A, Wang M, (2016) Multi-model seasonal forecast of Arctic
663 sea-ice. *Forecast uncertainty at pan-Arctic and regional scales. Clim Dyn* (in press) doi:10.1007/s00382-016-3388-9
- 664 Bosilovich MG, Akella S, Coy L, Cullather R, Draper C, Gelaro R, Kovach R, Liu Q, Molod A, Norris P, Wargan K,
665 Chao W, Reichle R, Takacs L, Vikhliayev Y, Bloom S, Collow A, Firth S, Labow G, Partyka G, Pawson S, Reale O,
666 Schubert SD, Suarez M (2016) MERRA-2. Initial evaluation of the climate. *Tech. Note NASA/TM–2015–104606*
667 Vol. 43 Koster RD, ed
- 668 Cai M, Kalnay E, Toth Z (2003) Bred vectors of the Zebiak-Cane model and their potential application to ENSO
669 predictions. *J Climate* 16:40–56 doi:10.1175/1520-0442(2003)016<0040:BVOTZC>2.0.CO;2
- 670 Cavalieri DJ, Parkinson CL, Gloersen P, Zwally HJ (1996) Sea ice concentrations from Nimbus-7 SMMR and
671 DMSP SSM/I-SSMIS passive microwave data, version 1. NASA National Snow and Ice Data Center Distributed
672 Active Archive Center, Boulder, Colorado doi:10.5067/8GQ8LZQVL0VL
- 673 Chang P, Ji L, Li H (1997) A decadal climate variation in the tropical Atlantic Ocean from thermodynamic air-sea
674 interactions. *Nature* 385:516 - 518 doi:10.1038/385516a0
- 675 Chou MD (1990) Parameterization for the absorption of solar radiation by O₂ and CO₂ with application to climate
676 studies. *J Climate* 3:209-217
- 677 Chou MD (1992) A solar radiation model for use in climate studies. *J AtmosSci* 49:762-772.
- 678 Chou MD, Suarez MJ, Liang XZ, Yan M (1994) An Efficient Thermal Infrared Radiation Parameterization for Use
679 in General Circulation Models. Technical Report Series on Global Modeling and Data Assimilation NASA/TM-
680 1994-104606 Vol. 3
- 681 Cullather RI, Bosilovich MG (2012) The energy budget of the polar atmosphere in MERRA. *J Climate* 25:5-24
682 doi:10.1175/2011JCLI4138.1
- 683 Deser C, Tomas R, Alexander M, Lawrence D (2010) The seasonal atmospheric response to projected Arctic sea ice
684 loss in the late twenty-first century. *J Climate* 23:333-351 doi:10.1175/2009JCLI3053.1
- 685 Eliassen A (1954) Provisional report on calculation of spatial covariance and autocorrelation of the pressure field.
686 Report 5. Videnskaps Akademi et Institut for Vaer Og Klimaforskning, Oslo, Norway, 12pp
- 687 Ferrari R, Nikurashin M (2010) Suppression of eddy diffusivity across jets in the Southern Ocean. *J Phys Oceanogr*
688 40:1501–1519. doi:10.1175/2010JPO4278.1
- 689 Fetterer F, Knowles K, Meier W, Savoie M (2016) Sea Ice Index, version 2. National Snow and Ice Data Center,
690 Boulder, Colorado. doi:10.7265/N5736NV7

691 Fox-Kemper B, Ferrari R, Hallberg R (2008) Parameterization of Mixed Layer Eddies. Part I: Theory and
 692 Diagnosis. *J PhysOceanogr* 38:1145–1165.doi: 10.1175/2007JPO3792.1
 693 Fox-Kemper B, Danabasoglu G, Ferrari R, Griffies SM, Hallberg RW, Holland MM, Maltrud ME, Peacock
 694 S,Samuels BL (2011) Parameterization of mixed layer eddies. III: Implementation and impact in global ocean
 695 climate simulations. *Ocean Model* 39(1-2):61-78
 696 Garcia R, Boville B (1994) “Downward Control” of the Mean Meridional Circulation and Temperature Distribution
 697 of the Polar Winter Stratosphere. *J AtmosSci* 51:2238–2245 doi: 10.1175/1520-
 698 0469(1994)051<2238:COTMMC>2.0.CO;2
 699 Gaspari G, Cohn SE (1999) Construction of correlation functions in two and three dimensions. *Quart J Roy Meteor*
 700 *Soc* 125B 723-757
 701 Gent PR, McWilliams JC (1990) Isopycnal mixing in ocean circulation models. *J PhysOceanogr* 20:150–
 702 155.doi:10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2
 703 Griffies SM, Harrison MJ, Pacanowski RC, Rosati A (2004) A Technical Guide to MOM4. GFDL Ocean group
 704 technical report No. 5
 705 Griffies SM (2012) Elements of the Modular Ocean Model (MOM): 2012 release. GFDL Ocean Group Tech. Rep.7,
 706 618 pp
 707 Guemas V, Blanchard-Wrigglesworth E, Chevallier M, Day JJ, Déqué M, Doblas-Reyes FJ, FučkarNS, Germe A,
 708 Hawkins E, Keeley S, Koenigk T, Salas y Mélia D, Tietsche S (2016) A review on Arctic sea-ice predictability and
 709 prediction on seasonal to decadal time-scales. *QuartJRoyMeteorolSoc* 142:546-561 doi:10.1002/qj.2401
 710 Ham YG, Rienecker MM, Schubert S, Marshak J, YehSW, YangSch (2012) The Decadal Modulation of Coupled
 711 Bred Vectors. *Geophys Res Lett* 39: L20712 doi: 10.1029/2012GL053719
 712 Ham YG, Rienecker MM, Suarez MJ et al. (2014a) Decadal prediction skill in the GEOS-5 forecast system. *ClimDyn*
 713 42:1 doi:10.1007/s00382-013-1858-x
 714 Ham YG, Schubert S, Vikhliav Y et al. (2014b) An assessment of the ENSO forecast skill of GEOS-5 system.
 715 *Clim Dyn*43:2415 doi:10.1007/s00382-014-2063-2
 716 Hamilton LC, Stroeve J (2016) 400 predictions. *The SEARCH Sea Ice Outlook 2008-2015*. *Polar Geogr* 39:274-287
 717 doi:10.1080/1088937X.2016.1234518
 718 Hill C, DeLuca C, Balaji V, Suarez M, da Silva A (2004) Architecture of the Earth System Modeling Framework.
 719 *ComputSciEng*, 6, 18–28
 720 Holland MM, BaileyDA, Vavrus S (2011) Inherent sea ice predictability in the rapidly changing Arctic environment
 721 of the Community Climate System Model, version 3. *ClimDynam* 36:1239-1253doi:10.1007/s00382-010-0792-4
 722 Hunke EC, Lipscomb WH (2008) CICE: The Los Alamos Sea Ice Model, Documentation and Software Manual,
 723 Version 4.0. Technical Report, Los Alamos National Laboratory.
 724 Hurrell J, Hack J, Shea D, Caron J, Rosinski J (2008) A New Sea Surface Temperature and Sea Ice Boundary
 725 Dataset for the Community Atmosphere Model. *J Climate* 21:5145–5153 doi: 10.1175/2008JCLI2292.1
 726 Ingleby B, Huddleston M (2007) Quality control of ocean temperature and salinity profiles - historical and real-time
 727 data. *J Marine Syst* 65:158–175 doi:158-175 10.1016/j.jmarsys.2005.11.019
 728 Keppenne, CL, Rienecker MM, Kovach RM, Vernieres G (2014) Background error covariance estimation using
 729 information from a single model trajectory with application to ocean data assimilation into the GEOS-5 coupled
 730 model. *NASA/TM-2014-104606* Vol. 34
 731 Kirtman B, Min D, Infanti J, Kinter J, Paolino D, Zhang Q, van den Dool H, Saha S, Mendez M, Becker E, Peng P,
 732 Tripp P, Huang J, DeWitt D, Tippet M, Barnston A, Li S, Rosati A, Schubert S, Rienecker M, Suarez M, Li Z,
 733 Marshak J, Lim Y, Tribbia J, Pegion K, Merryfield W, Denis B, Wood E (2014) The North American Multimodel
 734 Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. *B Am*
 735 *MeteorolSoc* 95:585–601 doi: 10.1175/BAMS-D-12-00050.1
 736 Koster RD, Suarez MJ, Ducharme A, Stieglitz M, Kumar P (2000) A catchment-based approach to modeling land
 737 surface processes in a general circulation model: 1. Model structure. *J Geophys Res* 105 (D20): 24 809–24 822
 738 Koster RD and Suarez MJ (2001) Soil moisture memory in climate models. *J Hydrometeor* 2(6): 558-570
 739 doi:10.1175/1525-7541(2001)002<0558:SMMICM>2.0.CO;2
 740 Large WG, McWilliams JC, Doney SC (1994) Oceanic vertical mixing: A review and a model with a nonlocal
 741 boundary layer parameterization. *Rev Geophys* 32(4): 363–403, doi:10.1029/94RG01872
 742 Large W, Danabasoglu G, McWilliams J, Gent P, Bryan F (2001) Equatorial Circulation of a Global Ocean Climate
 743 Model with Anisotropic Horizontal Viscosity. *J PhysOceanogr* 31:518–536 doi: 10.1175/1520-
 744 0485(2001)031<0518:ECOAGO>2.0.CO;2
 745 Large W, Danabasoglu G (2006) Attribution and Impacts of Upper-Ocean Biases in CCSM3. *J Climate*
 746 19:2325–2346 doi: 10.1175/JCLI3740.1

Lee HC, Rosati A, Spelman MJ (2006) Barotropic tidal mixing effects in a coupled climate model: Oceanic conditions in the Northern Atlantic. *Ocean Model* 11(3): 464–477
 Levitus S, Antonov JI, Boyer TP, Locarnini RA, Garcia HE, Mishonov AV (2009) Global ocean heat content 1955–2008 in light of recently revealed instrumentation problems. *Geophys Res Lett*, 36, L07608 doi:10.1029/2008GL037155
 Lin S (2004) A “Vertically Lagrangian” Finite-Volume Dynamical Core for Global Models. *Mon Weather Rev* 132:2293–2307 doi: 10.1175/1520-0493(2004)132<2293:AVLFDC>2.0.CO;2
 Locarnini RA, Mishonov AV, Antonov JI, Boyer TP, Garcia HE, Baranova OK, Zweng MM, Johnson DR, (2010) *World Ocean Atlas 2009, Volume 1: Temperature*. Levitus S, Ed. NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 pp.
 Lock A, Brown A, Bush M, Martin G, Smith R (2000) A New Boundary Layer Mixing Scheme. Part I: Scheme Description and Single-Column Model Tests. *Mon Weather Rev* 128:3187–3199 doi: 10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2
 Lorenz EN (1963) Deterministic, non-periodic flow. *J AtmosSci*, **20**(2): 130-141 doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2
 Lorenz EN (1993) *The Essence of Chaos*. University of Washington Press: Seattle, WA, 227 p.
 Louis JF, Geleyn J (1982) A short history of the PBL parameterization at ECMWF. *Proc. ECMWF Workshop on Planetary Boundary Layer Parameterization*, Reading, United Kingdom, ECMWF, 25-27 November 1981
 Maslanik J, Stroeve J (1999) Near-real-time DMSP SSMIS daily polar gridded sea ice concentrations, version 1. NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado doi:10.5067/U8C09DWVX9LM
 McPhaden M & Co-Authors (2010) The Global Tropical Moored Buoy Array. *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2)*, Venice, Italy, 21-25 September 2009, Hall J, Harrison DE, Stammer D, Eds, ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.61
 McFarlane N (1987) The Effect of Orographically Excited Gravity Wave Drag on the General Circulation of the Lower Stratosphere and Troposphere. *J AtmosSci* 44:1775–1800 doi: 10.1175/1520-0469(1987)044<1775:TEOOEG>2.0.CO;2
 Molodtsov A, Takacs L, Suarez M, Bacmeister J, Song IS, Eichmann A (2012) The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna. NASA Technical Report Series on Global Modeling and Data Assimilation, NASA TM—2012-104606, 28, 117 pp.
 Molodtsov A, Takacs L, Suarez M, Bacmeister J (2015) Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2. *Geosci Model Dev*, 8:1339-1356 doi:10.5194/gmd-8-1339-2015
 Moorthi S, Suarez M (1992) Relaxed Arakawa-Schubert. A Parameterization of Moist Convection for General Circulation Models. *Mon Weather Rev* 120:978–1002 doi: 10.1175/1520-0493(1992)120<0978:RASAPO>2.0.CO;2
 Murray RJ, Explicit Generation of Orthogonal Grids for Ocean Models. *J ComputPhys* 126:251-273. doi:10.1006/jcph.1996.0136
 Oke PR, Brassington GB, Griffin DA, Schiller A (2010) Ocean data assimilation: a case for ensemble optimal interpolation. *Australian Meteorological and Oceanographic Journal* 59 67-76
 Palmer TN and Anderson DLT (1993) Scientific assessment of the prospects for seasonal forecasting. A European perspective. *Tech. Rep. No. 70*, ECMWF, Reading, England, 34p.
 Palmer TN and Anderson DLT (1994) The prospects for seasonal forecasting. A review paper. *Quart J Roy Meteorol Soc* 120:755-793 doi:10.1002/qj.49712051802
 Pearson K (1896) *Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity and Panmixia*. Philosophical Transactions of the Royal Society of London. 187:253–318 doi:10.1098/rsta.1896.0007
 Perovich DK, Richter-Menge JA (2009) Loss of sea ice in the Arctic. *Ann Rev Marine Sci* 1:417-441 doi:10.1146/annurev.marine.010908.163805
 Reichle R, Koster R, De Lannoy G, Forman B, Liu Q, Mahanama S, Touré A (2011) Assessment and Enhancement of MERRA Land Surface Hydrology Estimates. *J Climate* 24:6322–6338 doi: 10.1175/JCLI-D-10-05033.1
 Reynolds R, Smith T, Liu C, Chelton D, Casey K, Schlax M (2007) Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *J Climate* 20:5473–5496 doi: 10.1175/2007JCLI1824.1
 Richter-Menge J, Walsh JE, Brigham LW, Francis JA, Holland M, Nghiem SV, Raye R, Woodgate R (2012) Seasonal to Decadal Predictions of Arctic Sea Ice. Challenges and Strategies. Polar Research Board, National Academies Press, Washington, DC, 80pp
 Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, Bosilovich MG, Schubert SD, Takacs L, Kim G-K, Bloom S, Chen J, Collins D, Conaty A, da Silva A, Gu W, Joiner J, Koster RD, Lucchesi R, Molodtsov A, Owens T, Pawson S, Pegion P, Redder CR, Reichle R, Robertson FR, Ruddick AG, Sienkiewicz M, Woollen J (2011)

803 MERRA—NASA's Modern-Era Retrospective Analysis for Research and Applications. J Climate 24:3624-3628.
804 doi: 10.1175/JCLI-D-11-00015.1
805 Saji NH, Goswami BN, Vinayachandran PN, Yamagata T(1999) A dipole mode in the tropical Indian Ocean. Nature
806 401(6751): 360-363 Simmons HL, Hallberg RW, Arbic BK (2004): Internal wave generation in a global baroclinic
807 tide model. Deep-Sea Res Pt II, 51(25-26):3043-3068. doi:10.1016/j.dsr2.2004.09.015
808 Shi L, Hendon HH, Alves O, Luo J, Balmaseda M, Anderson D(2012) How Predictable is the Indian Ocean Dipole?.
809 Mon Wea Rev 140:3867–3884 doi: 10.1175/MWR-D-12-00001.1.
810 Smith WHF, Sandwell DT (1997) Global seafloor topography from satellite altimetry and ship depth soundings.
811 Science 277:1957-1962 doi:10.1126/science.277.5334.1956
812 Stieglitz M, Ducharne A, Koster R, Suarez M (2001) The Impact of Detailed Snow Physics on the Simulation of
813 Snow Cover and Subsurface Thermodynamics at Continental Scales. J Hydrometeorol 2: 228–24 doi: 10.1175/1525-
814 7541(2001)002<0228:TIODSP>2.0.CO;2
815 Stroeve J, Hamilton LC, Bitz CM, Blanchard-Wrigglesworth E (2014) Predicting September sea ice. Ensemble skill
816 of the SEARCH Sea Ice Outlook 2008-2013. Geophys Res Lett 41:2411-2418 doi:10.1002/2014GL059388
817 Stroeve J, Blanchard-Wrigglesworth E, Guemas V, Howell S, Massonnet F, Tietsche S (2015) Improving predictions
818 of Arctic sea ice extent. Eos 96 doi:10.1029/2015EO031431
819 Tietsche S, Day JJ, Guemas V, Hurlin WJ, Keeley SPE, Matei D, Msadek R, Collins M, Hawkins E (2014) Seasonal
820 to interannual Arctic sea ice predictability in current global climate models. Geophys Res Lett 41:1035-1043
821 doi:10.1002/2013GL058755
822 Thomas K, Robinson DA, Bhatt U, Bitz C, Bosart LF, Bromwich DH, Deser C, Meier W (2014) Linkages Between
823 Arctic Warming and Mid-Latitude Weather Patterns. Summary of a Workshop. Polar Research Board, National
824 Academies Press, Washington, DC, 79 pp
825 Toth Z, Kalnay E (1993) Ensemble Forecasting at NMC: The Generation of Perturbations. B Am Meteorol Soc
826 74:2317–2330 doi: 10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2
827 Vaughan DG, Comiso JC, Allison I, Carrasco J, Kaser G, Kwok R and coauthors (2013) Observations. Cryosphere.
828 In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment*
829 *Report of the Intergovernmental Panel on Climate Change*. Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK,
830 Boschung J, Nauels A, Xia Y, Bex V, Midgley PM, eds. Cambridge University Press, Cambridge, UK and New
831 York, NY, USA, p. 317-382 doi:10.1017/CBO9781107415324.012
832 Vernieres G, Rienecker MM, Kovach R, Keppenne C (2012) The GEOS-iODAS: Description and Evaluation.
833 NASA Technical Report Series on Global Modeling and Data Assimilation, NASA TM—2012-104606, Vol. 30
834 Wajswicz R, (2007) Seasonal-to-Interannual Forecasting of Tropical Indian Ocean Sea Surface Temperature
835 Anomalies: Potential Predictability and Barriers. J Climate 20:3320–3343 doi: 10.1175/JCLI4162.1
836 Wan L, Bertino L, Zhu J (2010) Assimilating Altimetry Data into a HYCOM Model of the Pacific: Ensemble
837 Optimal Interpolation versus Ensemble Kalman Filter. J Atmos Ocean Tech 27 753-765
838 Weller RA, Anderson JL, Arribas A, Dickinson RE, Goddard L, Kalnay E, Kirtman B, Koster RD, Richman MB,
839 Saravanan R, Waliser D, Wang B (2010) Assessment of Intraseasonal to Interannual Climate Prediction and
840 Predictability. National Academies Press, Washington DC, 181p.
841 Yang S, Kalnay E, Cai M, Rienecker M, Yuan G, Toth Z (2006) ENSO Bred Vectors in Coupled Ocean–
842 Atmosphere General Circulation Models. J Climate 19:1422–1436 doi: 10.1175/JCLI3696.1
843 Yang S, Kalnay E, Cai M, Rienecker M, (2008) Bred Vectors and Tropical Pacific Forecast Errors in the NASA
844 Coupled General Circulation Model. Mon Weather Rev 136:1305–1326 doi:10.1175/2007MWR2118.1
845 Yang SC, Keppenne C, Rienecker MM, Kalnay E (2009) Application of Coupled Bred Vectors to Seasonal-to-
846 Interannual Forecasting and Ocean Data Assimilation. J Climate 22: 2850-2870 doi:10.1175/2008JCLI2427.1

847 List of Acronyms

848 ACC	Anomaly correlation coefficient
849 AGCM	Atmospheric general circulation model
850 AOGCM	Coupled atmosphere-ocean general circulation model
851 ARCUS	Arctic Research Consortium of the United States
852 CICE	Los Alamos National Laboratory Community Ice Code
853 CMAP	CPC Merged Analysis of Precipitation
854 CMIP5	Coupled Model Intercomparison Project phase 5
855 CPC	Climate Prediction Center

856	CTD	conductivity-temperature-depth
857	DMSF	Defense Meteorological Satellite Program
858	EnKF	ensemble Kalman filter
859	EnOI	Ensemble optimal interpolation
860	ENSO	El Niño/Southern Oscillation
861	ESMF	Earth System Modeling Framework
862	GDAC	Argo Global Data Assembly Center
863	GEOS-5	Goddard Earth Observing System model, version 5
864	GEOS-iODAS	Goddard Earth Observing System integrated Ocean Data Assimilation System
865	GEWEX	Global Energy and Water cycle Exchanges project
866	GMAO	NASA Global Modeling and Assimilation Office
867	GPCP	Global Precipitation Climatology Project
868	IDM	Indian Ocean Dipole Mode SST index
869	LSM	Land surface model
870	MERRA	Modern-Era Retrospective analysis for Research and Applications
871	MOM4	Modular Ocean Model, version 4
872	MSSS	Mean square skill score
873	NMME	North American Multi-Model Ensemble
874	NSIDC	National Snow and Ice Data Center
875	NOAA	National Oceanic and Atmospheric Administration
876	OI	Optimal interpolation
877	OOPC	Ocean Observations Panel for Climate
878	PIRATA	Prediction and Research Moored Array in the Atlantic program
879	RAMA	Research Moored Array for African-Asian-Australian Monsoon Analysis and Prediction
880	SAFE	Spatial approximation of forecast errors
881	SETIO	Southeastern Tropical Indian Ocean SST Index
882	SMMR	Scanning Multi-channel Microwave Radiometer
883	SSM/I	Special Scanning Microwave Imager
884	SSMIS	Special Scanning Microwave Imager/Sounder
885	SST	Sea surface temperature
886	T2M	2-m air temperature
887	TA	Tropical Atlantic index
888	TAO/TRITON	Tropical Atmosphere Ocean program Triangle Trans-Ocean Buoy Network
889	TASI	Tropical Atlantic SST Index
890	WTIO	Western Tropical Indian Ocean SST Index
891	XBT	eXpendable Bathythermographs